

TENT COOPERATION TRE Y

PCT

NOTIFICATION OF THE RECORDING
OF A CHANGE(PCT Rule 92bis.1 and
Administrative Instructions, Section 422)

From the INTERNATIONAL BUREAU

To:

BERGVALL EFTRING, Stina, Lena
Dr Ludwig Brann Patentbyrå AB
P.O. Box 17192
S-104 62 Stockholm
SUÈDE

Date of mailing (day/month/year) 23 October 2001 (23.10.01)	IMPORTANT NOTIFICATION
Applicant's or agent's file reference P04646PC00	
International application No. PCT/SE00/01079	International filing date (day/month/year) 26 May 2000 (26.05.00)

1. The following indications appeared on record concerning:
☒ the applicant ☐ the inventor ☐ the agent ☐ the common representative

Name and Address ACTIVE BIOTECH AB P.O. Box 724 S-220 07 Lund Sweden	State of Nationality SE	State of Residence SE
	Telephone No.	
	Facsimile No.	
	Teleprinter No.	

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:
☒ the person ☐ the name ☒ the address ☐ the nationality ☐ the residence

Name and Address SBL VACCIN AB Lundagatan 2 S-105 21 Stockholm Sweden	State of Nationality SE	State of Residence SE
	Telephone No.	
	Facsimile No.	
	Teleprinter No.	

3. Further observations, if necessary:

4. A copy of this notification has been sent to:
☒ the receiving Office ☐ the designated Offices concerned
☐ the International Searching Authority ☒ the elected Offices concerned
☒ the International Preliminary Examining Authority ☐ other:

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland	Authorized officer Céline Faust
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38

PATENT COOPERATION TREATY

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

Commissioner
 US Department of Commerce
 United States Patent and Trademark
 Office, PCT
 2011 South Clark Place Room
 CP2/5C24
 Arlington, VA 22202
 ETATS-UNIS D'AMERIQUE
 in its capacity as elected Office

Date of mailing (day/month/year) 20 February 2001 (20.02.01)	
International application No. PCT/SE00/01079	Applicant's or agent's file reference P04646PC00
International filing date (day/month/year) 26 May 2000 (26.05.00)	Priority date (day/month/year) 28 May 1999 (28.05.99)
Applicant FOLKESSON, Anders et al	

1. The designated Office is hereby notified of its election made:

☒ in the demand filed with the International Preliminary Examining Authority on:
20 December 2000 (20.12.00)

☐ in a notice effecting later election filed with the International Bureau on:

2. The election ☒ was
☐ was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No.: (41-22) 740.14.35	Authorized officer F. Baechler Telephone No.: (41-22) 338.83.38
--	--

REC'D 08 OCT 2001

WIPO

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

14

Applicant's or agent's file reference P04646PC00/HA	FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/SE00/01079	International filing date (day/month/year) 26.05.2000	Priority date (day/month/year) 28.05.1999
International Patent Classification (IPC) or national classification and IPC ₇ C 07 K 14/225, A 61 K 39/12		
Applicant ACTIVE BIOTECH AB et al		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 4 sheets, including this cover sheet.

☒ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consist of a total of 3 sheets.

3. This report contains indications relating to the following items:

- I ☒ Basis of the report
- II ☐ Priority
- III ☒ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- IV ☐ Lack of unity of invention
- V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI ☐ Certain documents cited
- VII ☐ Certain defects in the international application
- VIII ☐ Certain observations on the international application

Date of submission of the demand 20.12.2000	Date of completion of this report 24.09.2001
Name and mailing address of the IPEA/SE Patent- och registreringsverket Box 5055 S-102 42 STOCKHOLM Facsimile No. 08-667 72 88	Authorized officer Yvonne Siösteen/EÖ Telephone No. 08-782 25 00

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/SE00/01079

I. Basis of the report

1. With regard to the **elements** of the international application:*

- ☐ the international application as originally filed
- ☒ the description:
 pages 1-3 and 5-13, 67, as originally filed
 pages _____, filed with the demand
 pages 4, filed with the letter of 05.09.2001
- ☒ the claims:
 pages _____, as originally filed
 pages _____, as amended (together with any statement) under article 19
 pages _____, filed with the demand
 pages 15 and 16, filed with the letter of 05.09.2001
- ☒ the drawings:
 pages 1-4, as originally filed
 pages _____, filed with the demand
 pages _____, filed with the letter of _____
- ☒ the sequence listing part of the description:
 pages 14-66, as originally filed
 pages _____, filed with the demand
 pages _____, filed with the letter of _____

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language _____ which is:

- ☐ the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- ☐ contained in the international application in written form.
- ☐ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. ☐ The amendments have resulted in the cancellation of:

- ☐ the description, pages _____
- ☐ the claims, Nos. _____
- ☐ the drawings, sheet/fig _____

5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2 (c)).**

* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).

** Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

1. The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non obvious), or to be industrially applicable have not been examined in respect of:

☐ the entire international application,

☒ claims Nos. 10 and 11

because:

☒ the said international application, or the said claims Nos. 10 and 11

relate to the following subject matter which does not require an international preliminary examination (*specify*):

See PCT Rule 67.1.(iv).: Methods for treatment of the human or animal body by surgery or therapy, as well as diagnostic methods.

☐ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. _____
are so unclear that no meaningful opinion could be formed (*specify*):

☐ the claims, or said claims Nos. _____ are so inadequately supported
by the description that no meaningful opinion could be formed.

☐ no international search report has been established for said claims Nos. _____

2. A meaningful international preliminary examination cannot be carried out due to the failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions:

☐ the written form has not been furnished or does not comply with the standard.

☐ the computer readable form has not been furnished or does not comply with the standard.

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**1. Statement**

Novelty (N)	Claims	<u>1-9, 12-13</u>	YES
	Claims		NO
Inventive step (IS)	Claims	<u>1-9, 12-13</u>	YES
	Claims		NO
Industrial applicability (IA)	Claims	<u>1-9, 12-13</u>	YES
	Claims		NO

2. Citations and explanations (Rule 70.7)

The claimed invention relates to peptides encoded by nucleotide sequences from Salmonella for use in medicine. The sequences encode two fimbrial structures of Salmonella enterica subspecies I and Salmonella enterica subspecies I serovar Typhi. It has been found that SEQ ID NO 1 is specific to all Salmonella enterica subspecies I. The peptides and the nucleotide sequences encoding them are very suitable for use in diagnosis and in vaccine preparations.

It is already known from Folkesson et al., Abstracts of the General Meeting of the American Society for Microbiology 97, page 219 (1997) that the fimbrial genes safA, safB and safC from Salmonella have been identified.

The genes safA, safB and safC are encoded by parts of the nucleotide sequence SEQ ID NO 1 of Sequence listing No.1 of the present application.

The sequences have, however, not been disclosed in the prior art. Neither have the unique specificities of the genes been known before. Because of their specificity they are very useful in vaccine preparations.

Therefore the claims are considered to be novel and to involve an inventive step.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 00/01079

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: C07K 14/225, A61K 39/112

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: C07K, A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Abstracts of the 97th General Meeting of the American Society for Microbiology, May 4-8, 1997, Miami Beach, Florida, A. FOLKESSON et al: D-67. "Cloning and Characterization of Genes Encoding a Novel Salmonella spp. Adhesive Structure". See abstract.	1-4,6,8,10, 13-14
	--	
P,X	Molecular Microbiology, Volume 33, No 3, 1999, Anders Folkesson et al., "j Multiple insertions of fimbrial operons correlate with the evolution of Salmonella serovars responsible for human disease", page 612 - page 622, see entire document	1-14
	-- -----	

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

2 October 2000

Date of mailing of the international search report

23 -10- 2000

Name and mailing address of the ISA/

Swedish Patent Office

Box 5055, S-102 42 STOCKHOLM

Facsimile No. +46 8 666 02 86

Authorized officer

Henrik Nilsson/GH

Telephone No. +46 8 782 25 00

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE00/01079

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 10-11
because they relate to subject matter not required to be searched by this Authority, namely:
see next sheet *
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).:

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE00/01079

★

Claims 10-11 relate to methods of treatment of the human or animal body by surgery or by therapy/diagnostic methods practised on the human or animal body/ Rule. 39.1.(iv). Nevertheless, a search has been executed for these claims. The search has been based on the alleged effects of the compounds/compositions.



To
THE SWEDISH PATENT OFFICE
Stockholm

International patent application no PCT/SE00/01079
Our ref: P04646PC00/HA:sd

With reference to our telephone communication of September 13, 2001, we wish to file the enclosed page 16, comprising amended claim 14.

The amendment consists of a specification of the purpose of the diagnostic method wherein the primers or probes of the invention are used, by adding at the end of the claim the passage "for the purpose of detecting *Salmonella enterica* subspecies I". A basis for this amendment may be found in the description, e.g. on page 7, lines 20-22 and lines 27-29.

Stockholm, September 13, 2001
Active Biotech AB et al
by:

Harriet Allee

DR LUDWIG BRANN PATENTBYRÅ AB. Intellectual Property & Law Firm.

Stockholm Office: Box 171 92, S-104 62 Stockholm, Sweden. Visiting address: Maria Skolgata 83. Tel: +46 (0)8 429 10 00. (Non-Office hours: 429 10 99) Fax: +46 (0)8 429 10 70 (Patents). Fax: +46 (0)8 429 10 80 (Trade Marks). E-mail: brann@brann.se

Uppsala Office: Box 1344, S-751 43 Uppsala, Sweden. Visiting address: Drottninggatan 7. Tel: +46 (0)18 56 89 00. Fax: +46 (0)18 56 89 39.

Head Office and Seat of the Board: Stockholm. Reg. No. 556483-6219. Members of the Association of Swedish Patent Attorneys. Authorized representatives before the EPO and the OHIM

9. A vector vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I serovar Typhi, comprising a host in which a recombinant vector comprising a nucleic acid sequence selected from Sequence Listing No. 2 (SEQ ID NO 2), has been inserted and, optionally, a pharmaceutically acceptable carrier.

5

10. A method for protection against diseases caused by *Salmonella enterica* subspecies I, comprising administering a vaccine according to any of claims 4, 6, and 8.

10

11. A method for protection against diseases caused by *Salmonella enterica* subspecies I serovar Typhi, comprising administering a vaccine according to any of claims 5, 7, and 9.

15

12. Antibodies directed against a peptide encoded by a nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2), for use in a diagnostic method.

20

13. Peptide encoded by a nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2), for use in a diagnostic method.

25

14. Primers for, or probes that hybridize with, a nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2), for use in a diagnostic method for the purpose of detecting *Salmonella enterica* subspecies I.



To
THE SWEDISH PATENT OFFICE
Stockholm

International patent application no PCT/SE00/01079
Our ref: P04646PC00/HA:sd

With reference to our telephone communication of September 2, 2001, we wish to file the following pages, comprising amendments as specified herein below:

Amendments

In the description:

On page 4, under the heading "Sequence listing", the words " Sequence Listing No.1, herein referred to as" have been inserted before "SEQ ID NO 1", and the words "Sequence Listing No. 2, herein referred to as" have been inserted before "SEQ ID NO 2".

In the claims:

In new claims 1-4, 6, 8 and 12-14 the words "SEQ ID NO 1" have been put in parenthesis and the words "Sequence Listing No.1" have been inserted before the parenthesis.

In new claims 1-3, 5, 7, 9 and 12-14 the words "SEQ ID NO 2" have been put in parenthesis and the words "Sequence Listing No.2" have been inserted before the parenthesis.

Furthermore:

New claim 1

line 1: "protein" has been amended to "peptide";
lines 1-2: "the group consisting of " has been deleted;
line 2: "or parts thereof" has been deleted;

New claim 2

line 1: "the protein" has been amended to "a peptide";
line 2: "the group consisting of " has been deleted;
line 3: "or antigenic fragments thereof" has been deleted;

New claim 3

line 1: "the group consisting of " has been deleted;
line 2: "or parts thereof" has been deleted;

DR LUDWIG BRANN PATENTBYRÅ AB. Intellectual Property & Law Firm.

Stockholm Office: Box 171 92, S-104 62 Stockholm, Sweden. Visiting address: Maria Skolgata 83. Tel: +46 (0)8 429 10 00.

(Non-Office hours: 429 10 99) Fax: +46 (0)8 429 10 70 (Patents). Fax: +46 (0)8 429 10 80 (Trade Marks). E-mail: brann@brann.se

Uppsala Office: Box 1344, S-751 43 Uppsala, Sweden. Visiting address: Drottninggatan 7. Tel: +46 (0)18 56 89 00. Fax: +46 (0)18 56 89 39.

Head Office and Seat of the Board: Stockholm Reg. No. 556483-6210. Members of the Association of Swedish Patent Attorneys. Authorized representatives before the EPD and the OHIM.

New claim 4

- line 2: "the protein" has been amended to "a peptide";
 "or parts thereof" has been deleted;
 "the nucleotide sequence according to" has been amended to "a nucleotide sequence selected from";
- line 4: "the protein encoded by SEQ ID NO 1" has been amended to "said peptide";
 "or antigenic fragments thereof" has been deleted;

New claim 5

- line 2: "the protein, or parts thereof," has been amended to "a peptide";
- line 3: "the nucleotide sequence according to" has been amended to "a nucleotide sequence selected from";
- lines 3-4: "the protein encoded by SEQ ID NO 2" has been amended to "said peptide";
 "or antigenic fragments thereof" has been deleted;

New claim 6

- line 2: "SEQ ID NO 1, or parts thereof" has been amended to "a nucleic acid sequence selected from Sequence Listing No. 1 (SEQ ID NO 1)";

New claim 7

- lines 2-3: "SEQ ID NO 2, or parts thereof" has been amended to "a nucleic acid sequence selected from Sequence Listing No. 2 (SEQ ID NO 2)";

New claim 8

- line 3: "SEQ ID NO 1, or parts thereof" has been amended to "a nucleic acid sequence selected from Sequence Listing No. 1 (SEQ ID NO 1)";

New claim 9

- line 3: "SEQ ID NO 2, or parts thereof" has been amended to "a nucleic acid sequence selected from Sequence Listing No. 2 (SEQ ID NO 2)";

New claim 12

- line 1: "the protein" has been amended to "a peptide";
- line 2: "the group consisting of" has been deleted;
- lines 2-3: "or antigenic fragments thereof" has been deleted;

New claim 13

- line 1: "Protein" has been amended to "Peptide";
- line 1-2: "the group consisting of" has been deleted;
- line 2: "or parts thereof" has been deleted;

New claim 14

- line 2: "the group consisting of" has been deleted.

Basis for the amendments**The description**

The insertion on page 4 in the description is for a clarifying purpose only: it makes clear that by the wording "SEQ ID NO 1" (and 2 respectively) is meant "Sequence Listing No. 1" (and 2 respectively), conforming better to the established terminology within the field.

The claims

New claim 1 refers to the use of a peptide (instead of a protein)

A basis for the use of a peptide may be found e.g. on page 6, first paragraph, where a method of obtaining a peptide according to the invention is outlined, and on page 6, third paragraph, where medical uses of such peptide are mentioned.

Similar amendments (from protein to peptide) to claims 2-5 and 12 and 13 have the same basis in the description.

All other amendments to the claims relate to linguistic improvements.

Stockholm, September 4, 2001

Active Biotech AB et al

by:


Harriet Allee

of respective phage insert is represented by horizontal bars. Name and size of the phage inserts are indicated on the left side of the figure.

Figure 2: Schematic representation of the pTY52 cosmid comprising the *tcf*-operon (SEQ ID NO 2).

5 A *tcf* specific PCR fragment of 11105 bp was cloned into the Expand vector I cosmid (Roche). The insert is represented with a thick black line while vector sequences are represented with thin lines. Relevant restriction sites sequences are indicated. The position of the *tcf*-operon, i.e. *tcfA*, *B*, *C* and *D* (SEQ ID NO 2), is represented by a shaded arrow.

10 Figure 3: The phylogenetic distribution of the identified genes on the cs7 insert was investigated using the well defined SARC collection, see Example 1.

Figure 4: A 2 kb large internal *EcoR* I fragment was used as a probe in a Southern blot of the SARC collection, see Example 2.

15 Sequence listing

Sequence Listing No. 1, herein referred to as SEQ ID NO 1, —DNA sequence of the genes encoding the precursor of the *saf* fimbriae unit of *Salmonella enterica* subspecies I.

20 Sequence Listing No. 2, herein referred to as SEQ ID NO 2, —DNA sequence of the genes which encode the precursor of the *tcf* fimbriae unit of *Salmonella enterica* subspecies I serovar Typhi.

Deposit information

25 The phages carrying the inserted SEQ ID NO 1, i.e. phages clones B1, D1, F11 and N10 (see Figure 1) have been given the ECACC Accession numbers 99051922, 99051923, 99051924, and 99051925, respectively.

The cosmid carrying the inserted SEQ ID NO 2, i.e. cosmid pTY52 (see Figure 2) has been given the ECACC Accession number 99051926.

The depositions were made May 19, 1999.

30 Detailed description of the invention

The present invention is based on the finding that two fimbrial operons, the *saf* operon and the *tcf* operon, are specific for *Salmonella enterica* subspecies I bacteria. Due to their specificity they can be used to provide vaccines against
35 *Salmonella enterica* subspecies I as well as detection methods for *Salmonella enterica* subspecies I. The *saf* operon is specific for all *Salmonella enterica*

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 December 2000 (07.12.2000)

PCT

(10) International Publication Number
WO 00/73336 A1

(51) International Patent Classification⁷: C07K 14/225, A61K 39/112

(21) International Application Number: PCT/SE00/01079

(22) International Filing Date: 26 May 2000 (26.05.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
9901961-4 28 May 1999 (28.05.1999) SE

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): ACTIVE BIOTECH AB [SE/SE]; P.O. Box 724, S-220 07 Lund (SE).

Published:

- With international search report.
- Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): FOLKESSON, Anders [SE/SE]; Råsundavägen 169B, S-169 36 Solna (SE), NORMARK, Staffan [SE/SE]; Valhallavägen 126 2tr, S-114 41 Stockholm (SE), LÖFDAHL, Sven [SE/SE]; Adolfsbergsvägen 36, S-168 66 Bromma (SE).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(74) Agents: BERGVALL EFTRING, Stina, Lena et al.; Dr Ludwig Brann Patentbyrå AB, P.O. Box 17192, S-104 62 Stockholm (SE).



WO 00/73336 A1

(54) Title: FIMBRIAL PROTEINS

(57) Abstract: The present invention is based on the finding that two fimbrial structures are specific for *Salmonella enterica* sub-species 1 bacteria. Due to their specificity they can be used to provide vaccines against *Salmonella enterica* subspecies I as well as for detection of *Salmonella enterica* subspecies I.

FIMBRIAL PROTEINSField of the invention

The present invention is based on the finding that two fimbrial operons, the *saf* operon and the *tcf* operon, are specific for *Salmonella enterica* subspecies I bacteria and therefore have therapeutic use. Due to their specificity they can be used to provide vaccines against *Salmonella enterica* subspecies I as well as for detection of *Salmonella enterica* subspecies I. The *saf* operon is specific for all *Salmonella enterica* subspecies I bacteria and the *tcf* operon is specific for the serovar Typhi of *Salmonella enterica* subspecies I.

All or part of the DNA-sequences of the genes encoding these proteins can be used as active agents in a vaccine against diseases caused by the *Salmonella enterica* subspecies I bacterial strains or for detection of said bacterial strains.

The present invention also relates to methods of isolating these fimbrial proteins, to antibodies directed against these proteins, and to a vaccine composition comprising these proteins or antibodies directed against these proteins for use in the treatment of infections caused by the *Salmonella spp.* The fimbrial proteins according to the invention or antibodies directed against them can be used for detection of *Salmonella spp.* bacteria.

Background of the invention

The members of genus *Salmonella spp* colonize and infect a wide range of different organisms. Many cause gastroenteritis and enteric fever in humans and domesticated animals while others are not associated with human disease (Saylers et al, 1994). The genus has been divided into two species, *Salmonella bongori* and *Salmonella enterica* where *enterica* can be further subdivided into seven subspecies, designated I, II, IIIa, IIIb, IV, VI, and VII (Reeves et al, 1989). *Salmonella enterica* subspecies I are preferentially associated with warm-blooded animals. Over 99% of all clinical *Salmonella* isolates are strains belonging to this subspecies, including serovars Typhimurium and Enteritidis, which are the major causes of *Salmonella* induced gastroenteritis in humans, and Typhi, the human specific causative organism of typhoid fever, the most severe form of human salmonellosis (Popoff et al, 1992).

Salmonella enterica subspecies I consists of over 1300 different serovars and is preferentially associated with warm-blooded animals (Bäumler, 1997). Over 99% of all clinical *Salmonella* isolates are strains belonging to this subspecies,

including serovars Typhimurium and Enteritidis, which are the major causes of *Salmonella* induced gastroenteritis in humans, and Typhi, the human specific causative organism of typhoid fever, the most severe form of human salmonellosis (Popoff and Le Minor, 1992).

Today gastroenteritis and enteric fever can neither be prevented nor treated with good results. Typhoid fever is a substantial public health problem in developing countries. Each year 33 million people become ill and over 500 000 people die from this infection (American Institute of Medicine, 1986). Typhoid fever can be prevented by vaccination with attenuated bacteria, such as Ty21 and Vi vaccines and whole cell vaccines. Whole cell vaccines show a high incidence of side effects (Ashcroft et al, 1964, Yugoslav Typhoid commission, 1964). The vaccines consisting of attenuated strains of *Salmonella typhi* suffer from serious drawbacks. They must be administered as three or four spaced doses in order to stimulate protective immune responses (Levine et al, 1989). The treatment of *Salmonella typhi* with antibiotics is jeopardized since there are strains of *Salmonella typhi* that are resistant to chloramphenicol, ampicillin, and trimethoprim as well as ciprofloxacin (i.e. multidrug-resistant strains) (Rowe et al, 1997).

Accurate detection of *Salmonella enterica* subspecies I is today not possible. *Salmonella enterica* subspecies I can today only be detected by antibodies directed against surface proteins of *Salmonella enterica* subspecies I. The use of the sequences according to the invention makes it for the first time possible to rapidly and accurately determine the presence of *Salmonella enterica* subspecies I.

For many pathogenic bacteria, there is evidence that the filamentous surface protein structures called pili (fimbriae) are connected to the adhesion of the bacteria to the host cells. Pili proteins are very antigenic and are easily purified. Therefore pili preparations have been used as antigens for vaccination.

Summary of the invention

The invention relates to the objects as defined in the claims. The main object of the present invention is to provide two fimbrial proteins that are specific for *Salmonella enterica* subspecies I bacterial strains, the nucleotide sequences

encoding said proteins, as well as the corresponding amino acid sequences of for therapeutic and diagnostic use. Further are recombinant microorganisms provided, in which the nucleotide sequences according to the invention have been inserted.

5

An object of the present invention is to provide vaccine compositions for use in the treatment of *Salmonella enterica* infective strains, essentially pure Saf and Tcf fili protein of *Salmonella enterica* subspecies I and *Salmonella enterica* subspecies I serovar Typhi, respectively, as well as antibodies directed to these fili proteins.

10

A further object of the present invention is to provide the DNA sequences of the genes encoding the Saf and Tcf proteins. These sequences can be used for recombinant production of the proteins and for the preparations of vector vaccines against *Salmonella enterica* subspecies 1 and *Salmonella enterica* subspecies 1 serovar Typhi, respectively, as well as for diagnostic purposes.

15

Yet another object of the present invention to use purified Saf and Tcf protein from *Salmonella enterica* subspecies 1 bacteria for active or passive immunization of mammals, i.e. the proteins according to the invention can be comprised in a vaccine composition or be used to raise antibodies which can be comprised in a vaccine composition.

20

Finally, an object of the present invention is to provide a method for preventing or reducing the possibility of *Salmonella* infection of a mammal by administering the vaccines according to the invention. The invention may be more fully understood by reference to the following drawings and detailed description.

25

Brief description of the drawings

30

Figure 1.

Schematic representation of phage clones (named N10, D1, B1, F11) covering the entire cs7 insert of *Salmonella enterica* serovar Typhimurium strain SR χ 3181, i.e. comprising the *saf* fimbrial operon, i.e. *safA*, *B*, *C* and *D* (SEQ ID NO 1).

35

The clones were selected from partial *Eco* RI and *Bam*HI libraries in the Lambda Dash II vector. The cs7 insert is represented by a bold line. The extent

of respective phage insert is represented by horizontal bars. Name and size of the phage inserts are indicated on the left side of the figure.

Figure 2: Schematic representation of the pTY52 cosmid comprising the *tcf*-operon (SEQ ID NO 2).

5 A *tcf* specific PCR fragment of 11105 bp was cloned into the Expand vector I cosmid (Roche). The insert is represented with a thick black line while vector sequences are represented with thin lines. Relevant restriction sites sequences are indicated. The position of the *tcf*-operon, i.e. *tcfA*, *B*, *C* and *D* (SEQ ID NO 2), is represented by a shaded arrow.

10 Figure 3: The phylogenetic distribution of the identified genes on the cs7 insert was investigated using the well defined SARC collection, see Example 1.

Figure 4: A 2 kb large internal *EcoR* I fragment was used as a probe in a Southern blot of the SARC collection, see Example 2.

15 Sequence listing

Sequence Listing No. 1, herein referred to as SEQ ID NO 1, —DNA sequence of the genes encoding the precursor of the *saf* fimbriae unit of *Salmonella enterica* subspecies I.

20 Sequence Listing No. 2, herein referred to as SEQ ID NO 2, —DNA sequence of the genes which encode the precursor of the *tcf* fimbriae unit of *Salmonella enterica* subspecies I serovar Typhi.

Deposit information

25 The phages carrying the inserted SEQ ID NO 1, i.e. phages clones B1, D1, F11 and N10 (see Figure 1) have been given the ECACC Accession numbers 99051922, 99051923, 99051924, and 99051925, respectively.

The cosmids carrying the inserted SEQ ID NO 2, i.e. cosmids pTY52 (see Figure 2) has been given the ECACC Accession number 99051926.

The depositions were made May 19, 1999.

Detailed description of the invention

30 The present invention is based on the finding that two fimbrial operons, the *saf* operon and the *tcf* operon, are specific for *Salmonella enterica* subspecies I bacteria. Due to their specificity they can be used to provide vaccines against
35 *Salmonella enterica* subspecies I as well as detection methods for *Salmonella enterica* subspecies I. The *saf* operon is specific for all *Salmonella enterica*

subspecies 1 bacteria and the *tcf* operon is specific for the serovar typhi of *Salmonella enterica* subspecies 1, see Examples 1 & 2.

The main object of the invention relates to two fimbrial operons, the *saf* operon and the *tcf* operon, that are specific for *Salmonella enterica* subspecies 1 bacteria for therapeutic use.

Another object of the present invention is to provide vaccines against *Salmonella enterica* subspecies 1 induced gastroenteritis, enteric fever and typhoid fever.

A further object of the present invention is to provide methods to detect *Salmonella enterica* subspecies 1. The nucleotide sequences according to the invention are useful for constructing vectors for use as vaccines for insertion into attenuated bacteria in constructing a recombinant vaccine, for insertion into a viral vector in constructing a recombinant viral vaccine, or for direct inoculation as a nucleic acid vaccine. The pili proteins according to the invention, or antigenic fragments thereof, can be used for active immunization and antibodies directed against them can be used for passive immunization. All these applications of the sequences according to the invention are obtained by applying standard techniques known to the man ordinary skilled in the art.

Vaccines against *Salmonella enterica* subspecies 1.

The genes encoding the *saf* and *tcf* fimbrial structures, or fragments thereof, may be incorporated into a bacterial or viral vaccine comprising recombinant bacteria, virus or fungi which are engineered to produce one or more immunogenic epitopes of the *saf* or *tcf* fimbrial structures. In addition, the genes encoding the *saf* and *tcf* fimbrial structures, or part thereof, operatively linked to regulatory elements, can be introduced directly as a nucleic acid vaccine, to elicit a protective immune response.

The proteins or antigenic fragment thereof, deduced from the nucleic acid sequences of the present invention are useful alone or in conventional vaccine mixtures in the vaccine compositions according to the invention. The proteins could be produced by chemical synthesis or recombinant expression according to conventional methods.

The proteins and peptides according to the invention can be obtained by using a host organism transformed or transfected with an expression vector obtained by insertion of a gene according to the invention, or part thereof, into a vector in a conventional manner. The vector which is used to construct the expression
5 vector is not particularly limited, but specific examples include plasmids such as pET (Stratagen) and the like; and phages such as M13 (NEB), phage display libraries and the like. As expression regulatory sequence can among others T7 promoters and lac promoters be used.

10 An appropriate host to be transformed or transfected with the expression vector can be chosen among for example *E.-coli*, *Salmonella* or *Bacillus subtilus*. The transformed or transfected host is cultured and proliferated under suitable conditions.

15 After culturing, the peptides of the present invention may be purified by, for example, chromatography, precipitation, and/or density gradient centrifugation. The thus obtained peptides can be used as a vaccine or for the production of antibodies directed against said peptides, which can be used for passive immunization.

20

The purified preparation containing one or several proteins according to the invention, or parts thereof, is then formulated as a pharmaceutical composition, as for example a vaccine, or in a mixture with adjuvants. If
25 desired the proteins are fragmented by standard chemical or enzymatic techniques to produce antigenic segments.

In formulating the vaccine compositions with the peptide or protein, alone or in various combinations, the immunogen is adjusted to an appropriate concentration and formulated with any suitable vaccine adjuvant. The
30 immunogen may also be incorporated into liposomes, or conjugated to polysaccharides and/or other polymers for use in a vaccine formulation.

The different vaccines according to the present invention are administered to mammals in many different ways. These include intradermal, intramuscular,
35 intraperitoneal, intravenous, subcutaneous, oral, and intranasal routes of administration. The vaccine doses will differ depending on circumstances such

as body weight, interferences with other administered medicaments etc. The upper limit is not critical unless the dose shows toxicity.

The peptides and proteins of the present invention are also useful to produce
5 monoclonal or polyclonal antibodies for use in passive immunotherapy against *Salmonella enterica* subspecies 1. Human immunoglobulin is preferred.

Antisera is obtained from individuals immunized with proteins or peptides according to the invention. The immunoglobulin fraction is then enriched, for example by immunoaffinity or affinity chromatography. Antibodies raised in
10 a suitable mammal or in the patient to be treated, can subsequently be administered locally or topically, e.g. orally to the patient.

Detection of *Salmonella enterica* subspecies I in general.

The sequences according to the invention, or part thereof, or fragments
15 hybridizing therewith, as well as the proteins according to the invention, or part thereof, and antibodies directed to said proteins, or antigenic fragments thereof, can be used in molecular diagnostic assays for the detection of *Salmonell enterica* subspecies I.

20 Nucleic acids having the nucleotide sequence according to the invention, or any nucleotide sequence hybridizing therewith can be used as a probe in nucleic acid hybridization assays for the detection of *Salmonella spp* in various tissues and body fluids of patients. The hybridization assay may be of any type including; Southern blots, Northern blots, colony blots.

25 PCR technology is the most preferred technology for detection according to the invention of *Salmonella enterica* subspecies 1. Primers of at least one selected from the 5' end and one from the 3' end can be used in PCR and other known tests to rapidly identify the presence of *Salmonella enterica* subspecies 1. This is
30 according to conventional techniques.

The isolated and purified proteins and peptides of the invention can be used as diagnostics to measure an increase in serum titer of *Salmonella enterica*
35 subspecies I-specific antibody since they bind strongly to these antibodies. A serum test sample can be screened for *Salmonella enterica* subspecies I by methods such as for example ELISA.

The invention further comprises the use of antibodies directed against the *saf* and *tcf* fimbriae structures for quantitative or qualitative determinations of the pili proteins of the invention, or fractions thereof, in cells, tissues or body fluids.

5

Detection of *Salmonella enterica* subspecies I by using nucleic acid hybridization technology

10

Nucleic acid hybridization technology can also be used to detect *Salmonella enterica* subspecies I according to the invention. The nucleic acid probes chosen from parts of the sequences according to the invention can be either DNA or RNA. DNA sequences complementary to the sequences according to the invention can also be used. The binding of the probe to the target sequence, i.e. the hybridization, must not be perfect. Variations and mutations of the sequences according to the invention can be used as long as they hybridize good enough to detect *Salmonella enterica* subspecies I. The preferred length of the nucleic acid probes is about 10 to 400 nucleotides, most preferred not longer than 100 nucleotides.

15

The nucleotide probe is preferably chosen from the parts of the sequences that have the least variation. In the most preferred embodiments when screening for SEQ ID NO 1 (the *saf* operon, specific for *Salmonella enterica* subspecies I) a nucleotide probe or PCR primer selected from nucleotides 37 368-37 868 should be avoided since this region is hypervariable.

20

25 The nucleic acid probes according to the invention are prepared by any conventional method such as organic synthesis, recombinant techniques, or isolation from genomic DNA.

30 The nucleic acid probes of the invention are labeled in a conventional manner to signal hybridization to target nucleic acid from *Salmonella enterica* subspecies I. The labeling may comprise a radiolabel, an enzyme, a bacterial label, a fluorescent label, an antibody, an antigen, a latex particle, an electron dense compound, or a light scattering particle.

35

The probes may be provided in a lyophilized form, to be reconstituted in a buffer appropriate for hybridization, or the probes may already be present in

such a buffer. The buffer may contain a suitable hybridization enhancer, detergent, carrier DNA, and a compound to increase the specificity.

Any conventional hybridization assay technique, such as dot blot hybridization, Southern blotting, sandwich hybridization, displacement hybridization and the like, can be used.

The target analyte polynucleotide of a microorganism may be in various media, most often in a biological, or physiological specimen. In most cases it is preferred to subject the specimen containing the target polynucleotide to any conventional extraction, purification, and/or isolation before conducting the analysis.

The sample containing the target analyte nucleotide sequence must often be treated to convert the DNA to a single-stranded form, which may be accomplished by a variety of conventional techniques, such as thermal or chemical techniques.

The following examples describe the isolation and specificity of the sequences according to the invention.

EXAMPLE 1

Identification and characterization of the *saf* operon.

The present inventors found, upon investigation of a 7 kb chromosomal region on centisome 7 originally isolated from the *S. typhimurium* strain SR-11 κ 3181, a region that exhibits many of the traits that define a pathogenicity island. It has a lower G+C composition than the average composition of the *Salmonella* genome and includes many sequences related to different mobile genetic elements. The region is not present in *E. coli* K12, and the *Salmonella* specific DNA is inserted between the tRNA gene *aspV* and the stop codon of *yafV*, a hypothetical protein upstream of the *yafH* gene at 5 min in the *E. coli* chromosome. This *Salmonella* specific insert encodes proteins creating adhesive structures and other virulence factors. Sequencing revealed genes encoding a new fimbrial operon that they designated *Salmonella* Atypical Fimbriae (*saf*), due to its relatedness to a subgroup of adhesive structures forming thin atypical fimbriae or non-fimbrial adhesins.

The *saf* operon consists of four contiguous genes, *safA*, *safB*, *safC* and *safD* that encode fimbrial subunit, periplasmic chaperone, outer membrane usher protein and alternative fimbrial subunit, respectively. The genes *safA*, *B*, *C* and *safD* encode putative proteins of 166, 244, 836 and 156 amino acids, respectively. Analyzes of clinical *Salmonella* isolates showed that DNA of 195 out of 198 clinical isolates belonging to *S. enterica* subspecies I hybridized with *safB* and *safC*, i.e. these sequences are common to more than 99% of the known *Salmonella enterica* subspecies 1 bacteria. The inventors showed that 58% of these clinical isolates carry the *safA*, see Table 1.

Table 1. The prevalence of the *saf* genes in clinical *Salmonella* isolates.

Serovar	<i>safA</i>	<i>safB</i>	<i>safC</i>	# isolates
<i>S. adelaide</i>	-	+	+	1
<i>S. agona</i>	+	+	+	6
<i>S. anatum</i>	-	+	+	3
<i>S. bareilly</i>	+	+	+	3
<i>S. blockley</i>	+	+	+	3
<i>S. bovismorbificans</i>	-	+	+	5
<i>S. braenderup</i>	-	+	+	4
<i>S. brandenburg</i>	+	+	+	1
<i>S. bredeney</i>	+/-	+	+	15
<i>S. chester</i>	+	+	+	1
<i>S. colindale</i>	-	+	+	1
<i>S. derby</i>	-	+	+	1
<i>S. dublin</i>	-	+	+	1
<i>S. eastbourne</i>	+	+	+	2
<i>S. emek</i>	+	+	+	1
<i>S. enteritidis</i>	-	+	+	8
<i>S. give</i>	-	+	+	1
<i>S. goettingen</i>	+	+	+	1
<i>S. haardt</i>	-	+	+	1
<i>S. hadar</i>	+	+	+	16
<i>S. heidelberg</i>	-	+	+	1
<i>S. hittingfoss</i>	+	+	+	5
<i>S. infantis</i>	-/+	+	+	6
<i>S. java</i>	-	+	+	1
<i>S. javiana</i>	-	+	+	1
<i>S. kottbus</i>	-	+	+	1
<i>S. livingstone</i>	-	+	+	1
<i>S. london</i>	+	+	+	1
<i>S. maastricht</i>	+	+	+	2
<i>S. mbandaka</i>	-	-	-	3
<i>S. montevideo</i>	+	+	+	1
<i>S. muenster</i>	-	+	+	1
<i>S. newport</i>	+	+	+	2
<i>S. ohio</i>	+	+	+	1
<i>S. oranienburg</i>	+	+	+	2
<i>S. panama</i>	+	+	+	3
<i>S. potsdam</i>	+	+	+	1
<i>S. rissen</i>	-	-	-	1
<i>S. saarbrücken</i>	-	+	+	1
<i>S. saint paul</i>	+	+	+	3
<i>S. schwartzengrund</i>	-	+	+	1
<i>S. singapore</i>	+	+	+	1
<i>S. stanley</i>	+	+	+	5
<i>S. subsp I 4.5, 12:-</i>	+	+	+	2
<i>S. subsp I 4.5, 12:b:-</i>	-	+	+	1
<i>S. subsp I 4.5, 12:i:-</i>	+	+	+	1
<i>S. subsp I spont</i>	-	+	+	1
<i>S. tennessee</i>	+	+	+	2
<i>S. thompson</i>	-	+	+	1
<i>S. typhi</i>	-	+	+	1
<i>S. typhimurium</i>	+	+	+	27
<i>S. virchow</i>	+	+	+	7
<i>S. weltervreden</i>	-	+	+	1
<i>S. worthington</i>	-	-	-	2
<i>S. subsp III</i>	-	-	-	1

The phylogenetic distribution of the identified genes on the cs7 insert was investigated using the well defined SARC collection, which showed that the presence of the *safA*, *safB*, *safC* and *safD* genes is restricted to *S. enterica* subspecies I (Fig. 3). This region is hence the first subspecies I specific genetic region to be identified with a broad distribution within the subspecies. Since the serovars of subspecies I constitute over 99% of human salmonellosis and are preferentially associated with warm blooded animals, it implicates a role for the *saf* adhesive organelle in the colonization of these organisms.

EXAMPLE 2

Identification and characterization of the *tcf* operon.

The present inventors found that *Salmonella enterica* subspecies I serovar Typhi contains DNA encoding an additional fimbrial operon, the *tcf* operon, in the *sinR-pagN* intergenic region. Southern blot analysis revealed a markedly different restriction pattern in *S. enterica* serovar Typhi than the other subspecies I isolates, suggesting that the *saf-sin* region in serovar Typhi might carry additional DNA relative to serovar Typhimurium strains. A PCR reaction (using a kit from Roche) was therefore performed using a *sinR* (5'-GTA AAT CGC TTA GTC GCC-3') specific forward primer and a *pagN* (5'-TCA ACT CAA CCT TCA GCC-3') specific reverse primer.

This primer pair produced, as expected, a product of 2 kb in serovar Typhimurium from the SARC collection, while from serovar Typhi the product was 10 kb. Thus, the neighboring *sinR* and *pagN* genes in serovar Typhimurium strains are separated by approximately 8 kb in serovar Typhi.

The Typhi specific PCR product was purified, digested partially with *EcoRI* and sub-cloned into pUC18 forming a set of overlapping clones. Sequencing of the clones revealed a putative fimbrial operon designated *tcf* for Typhi Colonizing Factor. Four ORFs, *tcfA,B,C,D*, have been identified with putative proteins having significant homology to CooB (38% identical over 192 aa), CooA (37% identical over 170 aa), CooC (34% identical over 872 aa) and CooD (31% identical over 272 aa), respectively. The Coo proteins are involved in the biosynthesis of the CS1 colonizing factor antigens of enterotoxigenic *E.coli* (Fig. 4) (Froehlich et al., 1994). The peptide of the *tcfB* ORF is also homologous to the CblA major fimbrial subunit protein (45% identical over 154 aa) of the cable

10

15

References:

American Institute of Medicine. (1986) New vaccine development: establishing priorities. Washington, DC: National Academy Press.

5

Ashcroft, M. T., Ritchie, J. M., Nicholson, C. C. (1964) *Am. J. Hyg.* 79:196-206.

Levine, M. M., Taylor, D. N., Ferreccio, C. (1989) *Pediat. Infect. Dis. J.*, 8:374.

- 10 Popoff, M. Y. & Le Minor, L. (1992) *Antigenic formulas of the Salmonella serovars* (WHO Collaborating Center for Reference and Research on *Salmonella*, Institute Pasteur, Paris).

- 15 Reeves, M. W., Evins, G. M., Heiba, A. A., Plikaytis, B. D. & Farmer III, J. J. (1989) *J. Clin. Microbiol.* **27**, 313-320.

Rowe, B., Ward, L.R., and Threlfall, E.J. (1997) *Clinical Infectious Disease* 24:(Suppl 1) S106-9

- 20 Salyers, A. A. & Whitt, D. D. (1994) *Bacterial Pathogenesis: A molecular approach*. (ASM Press, Washington D.C.).

Yugoslav Typhoid Comission. (1964) *Bull. WHO* 30:623-30.

Claims:

- sub
B4
- 5
1. Peptide encoded by a nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2) for use in medicine.
 2. Antibodies directed against a peptide encoded by a nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2) for use in medicine.
 - 10 3. Nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2) for use in medicine.
 4. A vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I, comprising a peptide encoded by a nucleotide sequence selected from
15 Sequence Listing No. 1 (SEQ ID NO 1), or antibodies directed against said peptide and, optionally, a pharmaceutically acceptable carrier.
 5. A vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I serovar Typhi, comprising a peptide encoded by a nucleotide sequence
20 selected from Sequence Listing No. 2 (SEQ ID NO 2), or antibodies directed against said peptide and, optionally, a pharmaceutically acceptable carrier.
 6. A nucleic acid vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I, comprising a nucleic acid sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and, optionally, a pharmaceutically acceptable
25 carrier.
 7. A nucleic acid vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I serovar Typhi, comprising a nucleic acid sequence selected from Sequence Listing No. 2 (SEQ ID NO 2) and, optionally, a pharmaceutically acceptable carrier.
30
 8. A vector vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I, comprising a host in which a recombinant vector comprising a
35 nucleic acid sequence selected from Sequence Listing No. 1 (SEQ ID NO 1), has been inserted and, optionally, a pharmaceutically acceptable carrier.

16

Sub B4
ADP 34 AMDP
5 9. A vector vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I serovar Typhi, comprising a host in which a recombinant vector comprising a nucleic acid sequence selected from Sequence Listing No. 2 (SEQ ID NO 2), has been inserted and, optionally, a pharmaceutically acceptable carrier.

Sub A1
10 10. A method for protection against diseases caused by *Salmonella enterica* subspecies I, comprising administering a vaccine according to any of claims 4, 6, and 8.

10 11. A method for protection against diseases caused by *Salmonella enterica* subspecies I serovar Typhi, comprising administering a vaccine according to any of claims 5, 7, and 9.

Sub B5
15 12. Antibodies directed against a peptide encoded by a nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2), for use in a diagnostic method.

20 13. Peptide encoded by a nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2), for use in a diagnostic method.

25 14. Primers for, or probes that hybridize with, a nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2), for use in a diagnostic method for the purpose of detecting *Salmonella enterica* subspecies I.

Claims:

1. Peptide encoded by a nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2) for use in medicine.
2. Antibodies directed against a peptide encoded by a nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2) for use in medicine.
3. Nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2) for use in medicine.
4. A vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I, comprising a peptide encoded by a nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1), or antibodies directed against said peptide and, optionally, a pharmaceutically acceptable carrier.
5. A vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I serovar Typhi, comprising a peptide encoded by a nucleotide sequence selected from Sequence Listing No. 2 (SEQ ID NO 2), or antibodies directed against said peptide and, optionally, a pharmaceutically acceptable carrier.
6. A nucleic acid vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I, comprising a nucleic acid sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and, optionally, a pharmaceutically acceptable carrier.
7. A nucleic acid vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I serovar Typhi, comprising a nucleic acid sequence selected from Sequence Listing No. 2 (SEQ ID NO 2) and, optionally, a pharmaceutically acceptable carrier.
8. A vector vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I, comprising a host in which a recombinant vector comprising a nucleic acid sequence selected from Sequence Listing No. 1 (SEQ ID NO 1), has been inserted and, optionally, a pharmaceutically acceptable carrier.

9. A vector vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I serovar Typhi, comprising a host in which a recombinant vector comprising a nucleic acid sequence selected from Sequence Listing No. 2 (SEQ ID NO 2), has been inserted and, optionally, a pharmaceutically acceptable carrier.

10. A method for protection against diseases caused by *Salmonella enterica* subspecies I, comprising administering a vaccine according to any of claims 4, 6, and 8.

11. A method for protection against diseases caused by *Salmonella enterica* subspecies I serovar Typhi, comprising administering a vaccine according to any of claims 5, 7, and 9.

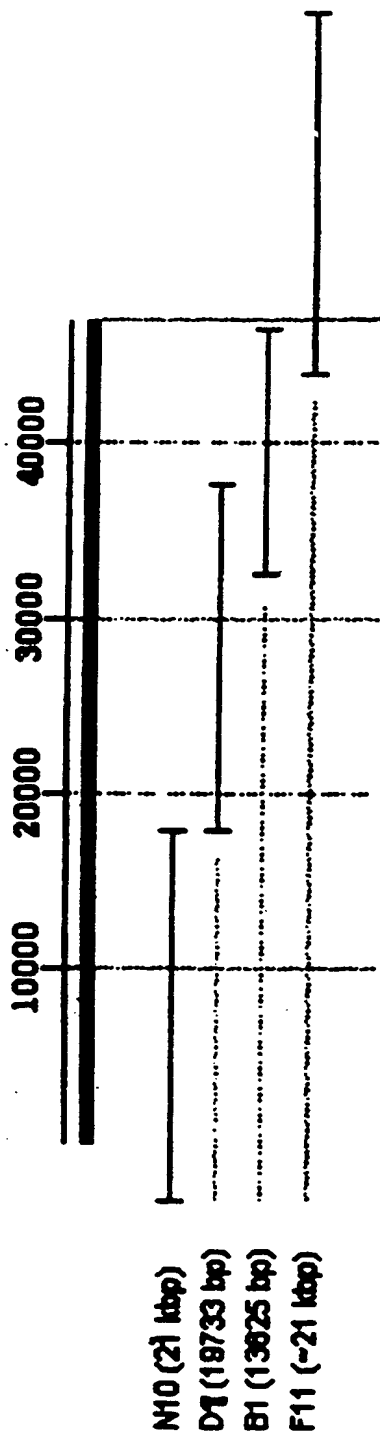
12. Antibodies directed against a peptide encoded by a nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2), for use in a diagnostic method.

13. Peptide encoded by a nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2), for use in a diagnostic method.

14. Primers for, or probes that hybridize with, a nucleotide sequence selected from Sequence Listing No. 1 (SEQ ID NO 1) and Sequence Listing No. 2 (SEQ ID NO 2), for use in a diagnostic method.

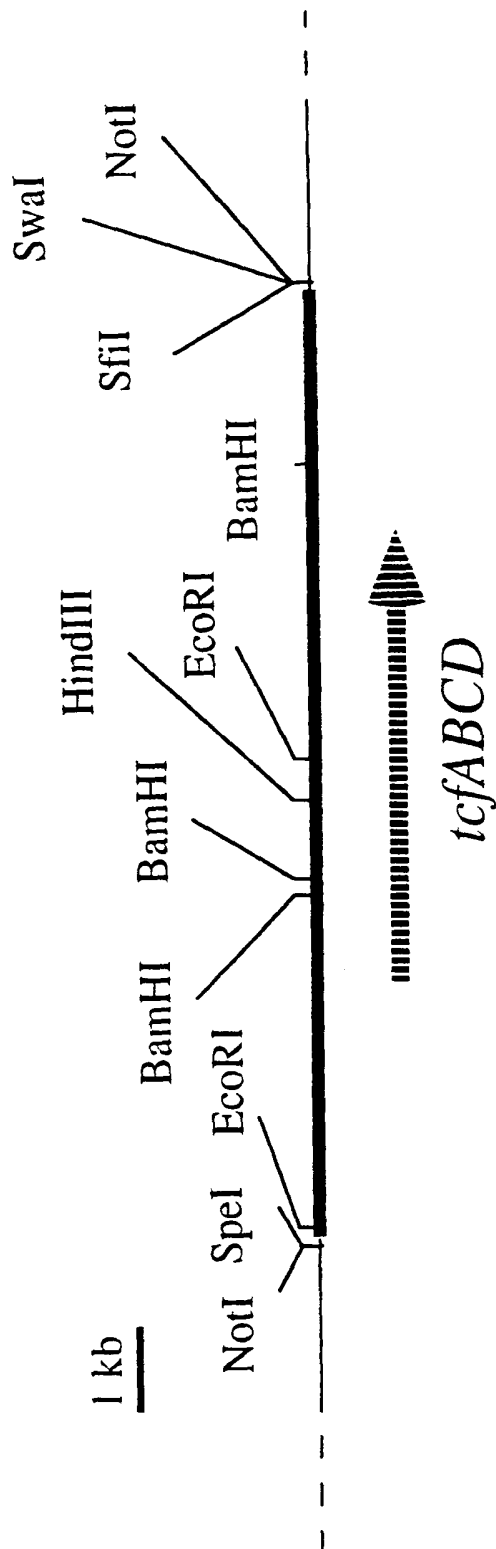
1/4

Fig. 1



2/4

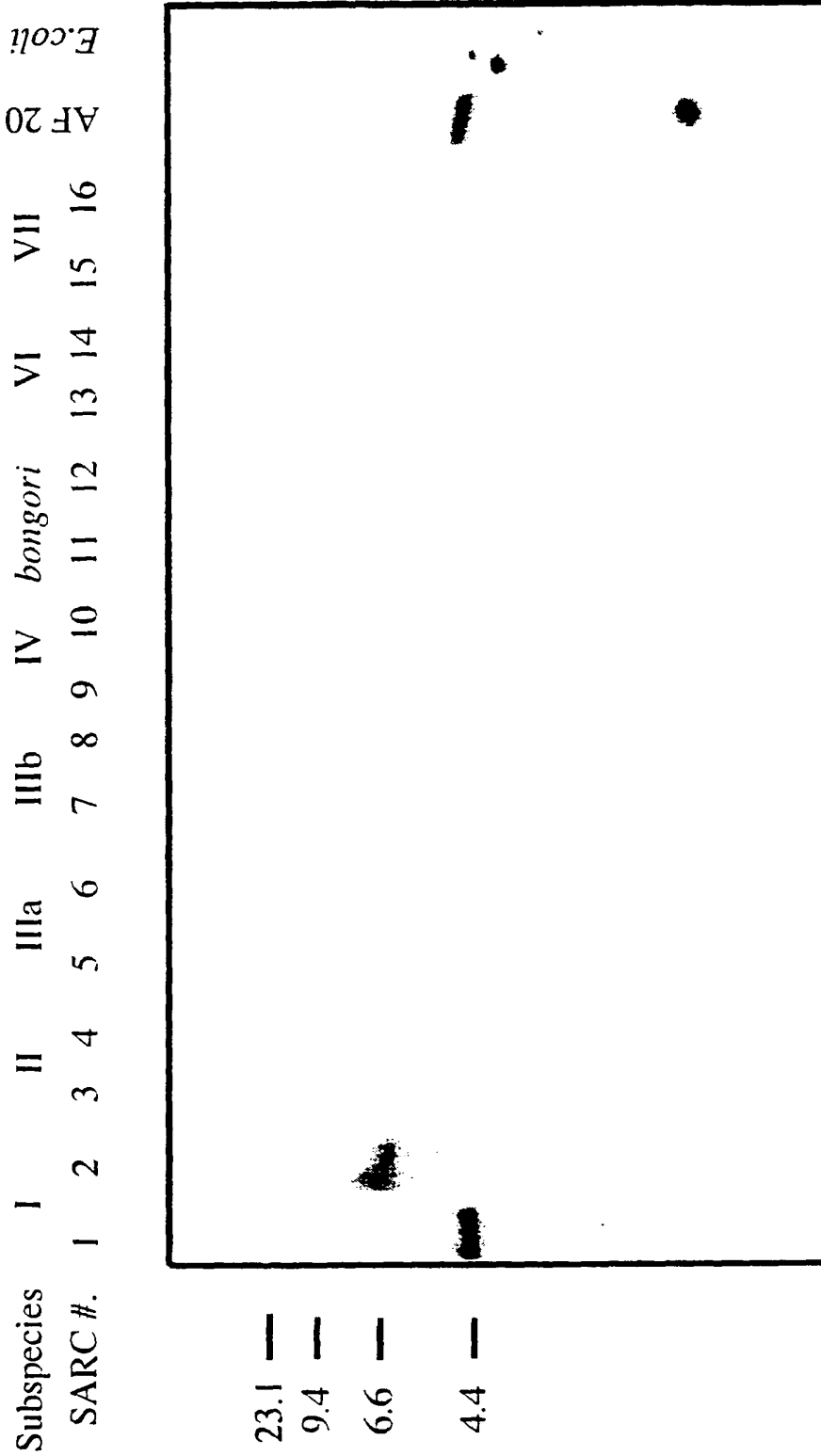
Fig. 2



3/4

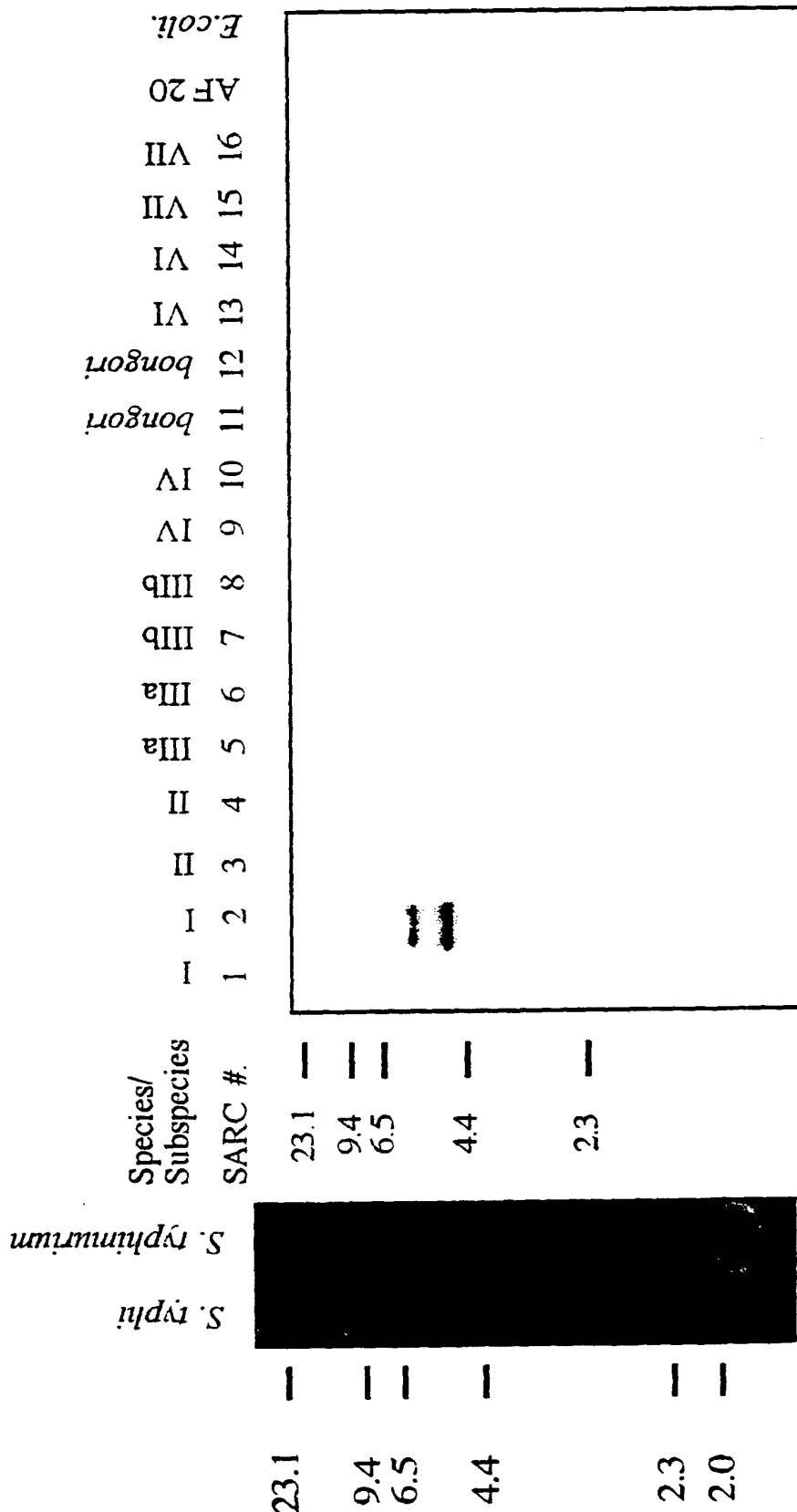
A

Fig. 3



4/4

Fig. 4



of respective phage insert is represented by horizontal bars. Name and size of the phage inserts are indicated on the left side of the figure.

Figure 2.

Schematic representation of the pTY52 cosmid comprising the *tcf*-operon (SEQ ID NO 2).

A *tcf* specific PCR fragment of 11105 bp was cloned into the Expand vector I cosmid (Roche). The insert is represented with a thick black line while vector sequences are represented with thin lines. Relevant restriction sites sequences are indicated. The position of the *tcf*-operon, i.e. *tcfA*, *B*, *C* and *D* (SEQ ID NO 2), is represented by a shaded arrow.

Figure 3.

The phylogenetic distribution of the identified genes on the cs7 insert was investigated using the well defined SARC collection, see Example 1.

Figure 4.

A 2 kb large internal *EcoR* I fragment was used as a probe in a Southern blot of the SARC collection, see Example 2.

Sequence listing

SEQ ID NO 1—DNA sequence of the genes encoding the precursor of the *saf* fimbriae unit of *Salmonella enterica* subspecies I.

SEQ ID NO 2—DNA sequence of the genes which encode the precursor of the *tcf* fimbriae unit of *Salmonella enterica* subspecies I serovar Typhi.

Deposit information

The phages carrying the inserted SEQ ID NO 1, i.e. phages clones B1, D1, F11 and N10 (see Figure 1) have been given the ECACC Accession numbers 99051922, 99051923, 99051924, and 99051925, respectively.

The cosmid carrying the inserted SEQ ID NO 2, i.e. cosmid pTY52 (see Figure 2) has been given the ECACC Accession number 99051926.

The depositions were made May 19, 1999.

Detailed description of the invention

The present invention is based on the finding that two fimbrial operons, the *saf* operon and the *tcf* operon, are specific for *Salmonella enterica* subspecies I bacteria. Due to their specificity they can be used to provide vaccines against *Salmonella enterica* subspecies I as well as detection methods for *Salmonella enterica* subspecies I. The *saf* operon is specific for all *Salmonella enterica*

Claims:

1. Protein encoded by a nucleotide sequence selected from the group consisting of SEQ ID NO 1 and SEQ ID NO 2, or parts thereof, for use in medicine.
2. Antibodies directed against the protein encoded by a nucleotide sequence selected from the group consisting of SEQ ID NO 1 and SEQ ID NO 2, or antigenic fragments thereof for use in medicine.
3. Nucleotide sequence selected from the group consisting of SEQ ID NO 1 and SEQ ID NO 2, or parts thereof, for use in medicine.
4. A vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I, comprising the protein, or parts thereof, encoded by the nucleotide sequence according to SEQ ID NO 1 or antibodies directed against the protein encoded by SEQ ID NO 1, or antigenic fragments thereof and, optionally, a pharmaceutically acceptable carrier.
5. A vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I serovar Typhi, comprising the protein, or parts thereof, encoded by the nucleotide sequence according to SEQ ID NO 2 or antibodies directed against the protein encoded by SEQ ID NO 2, or antigenic fragments thereof and, optionally, a pharmaceutically acceptable carrier.
6. A nucleic acid vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I, comprising SEQ ID NO 1, or parts thereof and, optionally, a pharmaceutically acceptable carrier.
7. A nucleic acid vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I serovar Typhi, comprising SEQ ID NO 2, or parts thereof and, optionally, a pharmaceutically acceptable carrier.
8. A vector vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I, comprising a host in which a recombinant vector comprising SEQ ID NO 1, or parts thereof, has been inserted and, optionally, a pharmaceutically acceptable carrier.

9. A vector vaccine for the protection against diseases caused by *Salmonella enterica* subspecies I serovar Typhi, comprising a host in which a recombinant vector comprising SEQ ID NO 2, or parts thereof, has been inserted and,
5 optionally, a pharmaceutically acceptable carrier.
10. A method for protection against diseases caused by *Salmonella enterica* subspecies I, comprising administering a vaccine according to any of claims 4, 6, and 8.
10
11. A method for protection against diseases caused by *Salmonella enterica* subspecies I serovar Typhi, comprising administering a vaccine according to any of claims 5, 7, and 9.
12. Antibodies directed against the protein encoded by a nucleotide sequence selected from the group consisting of SEQ ID NO 1 and SEQ ID NO 2, or antigenic fragments thereof, for use in a diagnostic method.
15
13. Protein encoded by a nucleotide sequence selected from the group consisting of SEQ ID NO 1 and SEQ ID NO 2, or parts thereof, for use in a diagnostic method.
20
14. Primers for or, probes that hybridize with a nucleotide sequence selected from the group consisting of SEQ ID NO 1 and SEQ ID NO 2, for use in a
25 diagnostic method.

aaccgtgcct gcgggggtctg gcactggcgc tggcgacgga gggatttccg gtgtggttgt 660
ctgtaccacg tcaggcagca gcgctaacaa ctgccgcagg cgtgaaaaat ctggagccag 720
atcgccaat gtctcacgcg ccacaccttg caggcgttct gcgccttccc gcgcttcgcg 780
aaacgccgcc agcggtaatg cccacggggc ttcgagggtcg gccagttgct ggcggacgga 840
ttcggggggc agcgcacag ccggggcggg agcggataac gcgcgctcca catcccgtac 900
ctgtaggcgc agggcgccac tgtttgacag cgtaataccg cgaatgtccg ccatcacgcc 960
ttcgtgatcc agcagcgccg ccaggggcatt actgcgcgcc agcgacgcac cttcaataga 1020
ttctgcggat gcgccttcgc cggtaagcag ctgtggatgg agtgcacag accagattac 1080
gctcagttct gccagttgtg tgagcatttc cgccagtccc tgtgcgccag cctgctggat 1140
gcgactgcgc agcagcagga tcaacaccgc gatatactta ctgcgcgtga gcagacggcg 1200
cgcatcacgt tcaatttccc gccagttcac ggcttccggc gtactgacaa aatcaccata 1260
ttgcgcctcc gcctgggggg cggcgcgggc gaacagcagc aggtattccc gatcgtactc 1320
cggatcgggg ccgcagggtt gttccgcgct taccggttta gtcagggaca tgtccatata 1380
attactctca gtgggttaag ccgtgttcag gttcaaaaat catgccggtc accggacggc 1440
catgcatggc ttacctgcc caggccgtcc agccaagccg ctgcgcgccg ccgatcaccg 1500
caggcggtgc ggcacgagc ttcaactgca gctcaatttc ccagcaatat taaaaaccga 1560
tgaacgtacg caccagttct gtcagcacgg gcagattgtt gccgccgggc agaaaacgca 1620
ggtaatcctc cagcgtgagc gggccgataa ttagccggaa tttgtactgc atatccggta 1680
cggcctggcc gataagcgcc ccgttgccca gcaccgaaga ctgcgcgggc gtgcccagac 1740
gggtgatttc atcgttccc accgttatcc agtgcagggc gaattctttc accgcaaaag 1800
gtacgctgaa atagtgcgc agcgtggcgg ccagcccgtc aggattgcgc gattcgcgta 1860
ccagatgggc ggaggctgcc aggcgaacat gatctgacag cgggctttcg gcgctttccc 1920
gtagatcctg ccgctgaga ctggcgatat aaaacgcaaa acggtcgtgt tccggtttgt 1980
ccagcccgcc accagcggac tgggcgctgc gccatgcctg ccagaactgc gtcagccagc 2040
ggcggtgaaa aatattggaa aaatgaacca gcgtcggatc gtgacgactc tctgagcggc 2100
tcagtgccag ctcggtatag tgcagcggca acgggccgtt tggcccccac agtccgaggc 2160
tgtacaggct cagggtcagg cgtccatcct gccagctgac ctgggcgatt tcccgtggcg 2220
caaaggctcat cgtcggcgtc tgtcccagac ggaatttttc catccgtggc ttgccagata 2280
cttgccctgcc ggagtatac agagctgggc atccacgcgc cgcacaggt tcaggaatcc 2340
gtaacgccag ggggttttct gcgcctgggt gagagcgcg gtcatacgt ccccgcttgc 2400

ccggttctga ccggccaggt catgacatgc ccgcgttgca tcgagtgcag cgtcatctgc 2460
gagaaggat taatggaaac atggcgggca atatagtgtt ccagcaccag accgaacagg 2520
taaggactga taccggaaaa tccttcttcg tccacgggtca gttcgcaact gacgccccgg 2580
ccatagacca acaggccgga gccgggcagg cggcgggtca ccgggggtgt tttgcagcca 2640
atcaggctgc gcacctggcg cgactgcggg ctgtcgtgag ccgggataaa cagattcagc 2700
agatcgcgca gcgcctggcc gccggtgcgg tgatccagat cggccagcgg cagataatta 2760
aacgacaact gccggatcag ccgccaggcc atttcgcgtt cagccagcgg cggctgcggc 2820
gggcgcggcg gtctgataag acccacgccc gccaccggaa tcgccgcac tacggtcaga 2880
tcatccccgc cattacgtgg aataaggcag ggcagatcgc ggttagtcac cattgccgtg 2940
acgggtgatat ggcgcagatt ttccgggtag ggcgcttcat gctgatcaac cagcgagagg 3000
aagacttccg agccggtata ggggggttcgg gtgccatagc ggccggcggt ttctgacgag 3060
cggcgcggtt cacgacgcag tgaaaaataa cggccgtggt tgccttcgtc attattacgg 3120
gtgtgataca gcggacgaaa aatcatcttc cgtgtggttt ccgcttcag gccttcgact 3180
tcctgaacag aaaacacctc gtaatccagc ggacgggtac gatccaccac cagatgctgt 3240
tccgtcacgc tgtgagtgc ttcaatccgg gtggtggtgc gaggaagcag gttgatcacc 3300
ggcgtacaga acaggctgaa ctgtgcagcg tccgtctgat gaatcagcca gtccggcggc 3360
aggcgggtta gcatgatgac aatttccgcc acattgcctt gcaccttttg taaccgggca 3420
gacaatccgg tcgggggtgaa gaagtaaaac cgttccggac aggcgaaaaa ttcattgcagc 3480
agattatggc cgtgaaacac gttccaggcg agcggtagca gcccttgcct tggctccagc 3540
cttcgtgcgc ccaccgggtg ttgaagattc acattcagtt cggcgtcaaa gtgaccgggt 3600
tcaccggcca gtgtggcgac ggcgctggtg ttagcagct caaacagggt tgacgcaatg 3660
cgttcttcgc cgcagaggta aaagggcagc cgtgccggac cggccagctc gctgaaagtc 3720
agttccccga aggttcgcag ggtgatgcgc aatgccccgg cgacatgaat attaggcggc 3780
agatagcggc gcaggcggg catatccggc ggcgcggcgg tcaggcgtac ctctcgtatg 3840
gacagcggcc acagcgtgac gtcctggctg ctgcgaaact ggcaggcggc attttcgcct 3900
tcgggatgg gggaaacgaa cgcggtatcg cgcggcacgg tgacctttt cggcaggctg 3960
cttctcgtcg tatcgggata cagctttacc actgccatcg atggcgtggg ggtgacgtaa 4020
ttggggctga cgacttccag taaccgctgt gtgaagcggg gaaactcggc gtcaattttt 4080
agctgagtgc gggcgctcag aaagctgaac gcctcgatca tgcgttcac atacgggtcg 4140
gcaatatcgg ttccctgcat cccagtcgg gcggcaattt tgggatggag ggtggcgaac 4200
tcagcaccgg tctcccgag gtagctcagt tcgcggtgt aatactccag tagccgtgga 4260

gggttctgcg gattgagctg ttgttgttta aaccagtcca gatccaggta ctcagcaacg 6120
 gattcagatg actgaaccaa ctgaggaaac gaaacatcgg gacgggactc atcattgggtg 6180
 cgggcaaggc accaggatga cagaccccat tcaatcgtat ctccatcatt cagccgcata 6240
 cggtgatgag gttccatcac ctgctcattg acggcacagc acagtccatc agagtgatta 6300
 acaatccacc acgcctcttc gtggcgaaac agactgagcg cgatatcgtc tttcccatgt 6360
 tgctcaaaaa gcggaatgta gggggcgggc gcagtaaata cgatgtatgt ccctgcggga 6420
 tactgtgcgt tatccgttat accctgcgac tgtattttcc gtaattccca catgtcttgt 6480
 ttcattatgt cccttaaatg tatttatctt tctggaggaa acgtttaggg agttttaatt 6540
 cataaataat tttactaaa tttataggga gtcattattg atgacaccct ttttattatt 6600
 ttcgtcgtga atgcattgtt gtatgcatag atgtctttct tgaaatatta tttcttttaa 6660
 ttctgcaatt gcgattttta tcttgctgtt cattgattat ttaatatata cggcttagtt 6720
 gctaattagt ccctgatctg tatcattgtt ttgtttcgat attttttcga ggctatcaat 6780
 aaagaattgt ttatatattt atatgcattg atgcattatt tttatgaatt tttatgtcac 6840
 aaggcataac acatggaaac tctgtttca cgcagtgcgt tgtatggaaa actggccggc 6900
 ccactattcc ggctcgtgga atcggcaacg gcattttgca aactacgctc taatccctgg 6960
 gttgagctga ctactggct gcaccagtta acacagcagc ccgataacga tattctccac 7020
 gttcttcggc attaccagat ccctctttct gatgtggaga aagcgttact ccggcaactg 7080
 gatatgctgc ccgcggggc cagcgccatt agtgattttt ctcaccatat cgatctcagc 7140
 gttgaaaagg cctggatgct ggcgagcgtc cgttacggcg ataacaaaat tcgcagcggc 7200
 tggttgctgc tggccttgtt gaccacgcca gaactgcgtc ggggtactgag cagtatctgc 7260
 gcgcgctgg ccacgcttcc ggttgatgaa ctgacggaaa tactgccctc gttgatcgaa 7320
 acatcgccgg aagcgcagga gcgcccctac gacggctccg gcctggcatc agccattccc 7380
 ggtgaaagca gtcaggcgat tcccaacggc gggcaggacg gtaaattccg gctggcaaaa 7440
 tactgtcagg acatgacggc acaggcgcgc gacggcaaaa tcgaccgggt gacggggcgt 7500
 gagcatgaaa tccgcaccat gacggatatt ctgctgcgcc gtcgccagaa taatccacta 7560
 ctgactgggt aggcggggcgt cgggaaaacg gcggctcgtc aagggttttg cctcgcgatt 7620
 gcgcaggggg aagtgcgcgc cgcgctgcgg gaagtacggc tactggcgct ggacgttggc 7680
 gctctgttgg ccggagccag catgaaaggc gagtttgaat cgcgtctgaa agggttactg 7740
 gaagaggccg ggcgctcgcc gcagccgggtt attctgtttg ttgatgaagt tcacactctg 7800
 gtgggcgcgg gcggcgcatc cggcacgggc gatgccgcta acctgctgaa accggcgctg 7860
 gcgcgcggca ccctgcggac tatcggcgcc accacctgga gcgaatacaa gcgccatatt 7920

gagaaagatc	cggcgctgac	ccgtcggtttt	caggtgttgc	agattgccga	accggaagag	7980
atccccgcaa	tggaaatggt	gcgtggtctg	gtggatacgc	tggaaaaaca	ccataacgta	8040
ctgattctgg	atgaggcggt	acgtgcggcg	gtacagcttt	ctcaccgcta	cattccccgc	8100
cggcagttgc	cggataaggc	catcagcctg	ctggataccg	ccgcggccccg	cgtggcgctg	8160
acgctgcaca	cgccgcctgc	cagcgtacag	ttcctgcgcc	agcagctaaa	agcggcggaa	8220
atggaacggt	cgctgttgca	gaagcaggaa	aaaatgggga	ttcagtcaga	tgagcggcgc	8280
gatgcgctga	tggcgcgaat	tttctcgctc	aacaatgaac	tgactgcatc	cgaatccccg	8340
tggcagcggg	agctggaact	ggtacatacg	ttgcaggaac	tgcgtctcgc	agagtctgat	8400
gctgatgaca	aaaccacgct	gcaacaggcc	gaaacggcgc	taagggagtg	gcagggcgac	8460
gcgccggtgg	tgttccccga	agtcagcgcg	gcggttgtcg	cggcgattgt	cgccgactgg	8520
accggtattc	ctgctggggcg	catggtgaaa	gatgaggcca	gccagggtgct	ggaactgcct	8580
gcccgactgg	cgcaacgcgt	taccgggcaa	gacggcgcgc	tggcgcagat	tggtgaaact	8640
attcagaccg	ccagggcggg	actgggcgat	ccacgcaaac	cgggtgggcgt	gtttatgctg	8700
gccggggccgt	ccggtgtcgg	taaaaccgaa	accgcgctgg	cgctggcgga	ggctatctac	8760
ggcggcgagc	agaacctggt	aaccatcaat	atgagcgagt	tccaggaggc	tcacaccgtt	8820
tccacgctga	aaggcgcgcc	gcccggtat	gtgggctatg	gcgagggtgg	tgtgctgacg	8880
gaagcggtgc	gtcgccaccc	ctggagcgta	gtgctgctcg	acgagatcga	aaaagcgcac	8940
catgacgtcc	acgaactctt	ctatcagggt	tttgacaagg	gtgggatgga	ggacggcgag	9000
ggtacacatg	tcgatttcaa	aaacaccacg	ctattactca	ccaccaacgt	gggttccgac	9060
ctcatcagcc	agatgtgtga	agatccggcc	ttaatgcccg	atgctacggg	gcttaaagag	9120
gcgctaattg	cggaattgcg	caagcatttc	ccggcggcac	ttctggggccg	cgtgacggtg	9180
atcccttacc	tgccgctgga	cgaaacgtcg	cgtggcgctga	ttgcccgtct	gcaccttgac	9240
cggctggtgg	cgcggatgag	tgaacagcac	ggcgtagcgc	tgacgtatag	cgaggaactg	9300
gtcgcacata	ttgtggcgctg	ctgtccaatg	catgaaacgg	gcgcgcgggt	gctgattggc	9360
tacatcgaac	agcacattct	gccacgactg	tcgcgctact	ggttgcaggc	catgacggaa	9420
aaagccgcga	tcaggcagat	tgatatcggc	gttaatgggtg	atgagcagat	tgtttttgag	9480
atcgtttgct	gaaaccggcc	gttcgaagtg	tccgtagtgc	gatttttaaaa	actgtaccgg	9540
tataccgctc	cccttgcggc	aaccagttga	ctaaaaagga	aatgaaggat	tatggctatc	9600
aacaatagcg	cgcagaaaatt	catcgcgcgc	aaccgcgcgc	cgcgcgtgca	gattgaatat	9660
gacgtagaga	tttacggttc	cgagaaaaaa	atcgagctgc	cgttcgtgat	ggcgggtgctg	9720

gccgatctgg ccgggaaacc gcgtgaagaa ctgccgccgg tgacggatcg caaatctctc 9780
 gatattgata ttgataactt caatgagcgc atgaaagcca ttgcgccgcg cgtggcggttc 9840
 gctgtgccga atacgctgac ggggtgaaggc cagttgatgg tcgatatcac gctggaaaaat 9900
 atggacgact ttctgccggc acagattgcc cgcaagggtgg acgccctgaa ccagttactg 9960
 gaagcccgca ctcaactggc gaacctccag acctacatgg atggcaaggc gggggscggaa 10020
 aatctgggtca ataaactggt gcaggaccgc actctgctga aaacgctggc gaatgcgcgcg 10080
 aaatcgggcg ctaccagca agatgtgtca gcggataatg aatcagcgga ataacgtcga 10140
 atttttaagg aattttcatg gcaaacagta atatgcaggc aaccgacgcg gttgctcagg 10200
 ataccgcctc cgcattccgt gaatttgatg cgttgctgaa tcaggccttc cgaccaaga 10260
 ctaccagggc ggcaaaagcc gtggaagccg cgggtgcagac gctggcggaac acgatcaccg 10320
 tcagcgatga cgctataaaa agcatcagcg ctattattgc gcagatcgac tttaaactga 10380
 ccgaacagat caaactgac ctgcaacatc ccgactggca gaagctggaa tctctgtggc 10440
 gcggtatgga gcatctgggt tacaacaccg agaccgacga aaagctgaaa attcgcttca 10500
 tgaatctgtc aaaagatgaa ttgcggcgca acatgaagcg ttacaagggc atcgcttggg 10560
 atcaaagccc gatgttcaag aaactgtatg aagccgaata cggccagtta ggtggcggaac 10620
 cttatggctg tatcattgcg gattactact tcgaccatac accgcccgat gtggatctgc 10680
 ttggtcttat cgccaaagtc gccgcgtcgg cccatgcgcc gtttattgcc ggggcttccc 10740
 cctcggtact gcaaatggac tcctggcagg aactggcgaa tccccgcgac ctgacaaaaa 10800
 tcgtcaccca gaacctggaa tatgcgccgt ggaactcgtc gcgggctagc gaagactccc 10860
 gttatatggc cctgacgat ccgcgttttc ttgccgcct gccgtatggc gcaaaaacca 10920
 acccggtgga cgagtttgat tttgaagaag atgcggatgg ttctgaccat accaaatacg 10980
 tctggagcaa cgcgccctac gcgatggcg taaacatcaa ccgttccttc aaacactacg 11040
 gctgggtgtac gttgattcgc ggtgtggaat caggcggtgc ggtggaaaaat cttccctgcc 11100
 ataccttccc gactgacgat gggcgcggtg acatgaaatg cccgaccgaa atcgccatct 11160
 ctgaccgccc cgaggctgaa ctggcgaaaa acggttttat cccgttgatc caccgtaaaa 11220
 actcagacta tgccgccttt atcggcgcac agtcgctgca aaaaccacag gaatactacg 11280
 atccggacgc gacggccaac gctaacctgt ctgcccgctc accgtacctg ttcgcctgct 11340
 cgcgcttcgc tcaattcctc aaatgtatcg tccgcgacaa aatcggttcc tttaaagagc 11400
 gtgaggatat gcagcgctgg ctaaatgaat ggattatgaa ttatgtcgac gccgatccgg 11460
 tgaactcctc gcaagaaact aaagcccgtc gtccgctggc tgccgctgaa gtagtggtgg 11520
 aagaggctga aggcaatcca gggtattacg acgcgaaatt cttcctgcgt ccgcatttcc 11580

agcttgaagg gctgacggga tcgctgcgcc tggtgacaaa actgccgtca gtgaagcagg 11640
 gcaatgcctg atatatatatt tgtgaatggt taagcgagtg aagtcagaga agatagagaa 11700
 tataaagagg gatatgaaga aaagaatttc gtctcgccca cggctctcgta aagggtgggg 11760
 acgtaatgat gacacatatc cgaatgccag taacaatgcc gaagcttttt atattcattga 11820
 gtaggaaata catattatga ccataagccc aacttttcat ctgttacctg gtattgttct 11880
 gctcttttca caatatgctg tagcctggga agtcagttgc cgggctgcta ttgataactca 11940
 gtcttctgct gtgagcctga agtctgatgt cccagcggcg tggcagcttt cttcccgcata 12000
 tatgtcgcgt ttatgggttaa gtagtattgg ggtaacgcag ggtaaacctg aaaacctgat 12060
 ggatctcaaa ccagagacta aaaaagttaa cggtgaaaat tggctctgtat gggaaacaga 12120
 acgtggtagc gataaagaaa ccgatcgcta ttgggtttcg tgtatttatg gtcattgaaca 12180
 gatatgggtg acgcaaccaa tacctgcttc tctactcgc tgtaagactc gtaattttga 12240
 gggatcgcca gaagaccagt ctgtatcttt tatctgtaat tagcgatttg agacgtgaaa 12300
 atttcagrac aggttatggt ttttattatc ggaagttagt aagcattatt tatatgcatt 12360
 aaataatgca aattcataaa ataactaaat acattatcgg taccggaaaa atatacagtc 12420
 ctctgttctc ctgaagttag tggagaagga ttctgtacgg caatgattta tctataaaca 12480
 aaaagatata gataaaatca ggtttatttt aagtaaaact taataaggat ataaaaatgg 12540
 cttatgacat ttttttgaaa attgacggca ttgatggcga gtcaatggat gacaaacaca 12600
 aaaatgaaat tgaagtactg agctggcgct ggaatattca tcaggaatcc accatgcacg 12660
 ccggtagcgg cctcggctcc ggtaaggctt ccgtcaccaa cctggatttt gatcactata 12720
 tcgaccgcgc cagcccgaac ctgttcaaat actgcgcctc cggcaagcac attccgcagg 12780
 ccattctggt tatgcgtaag gctggcggca atccgctgga gtacctcaag tataccttca 12840
 ccgacctgat tgtcgccgtg gtttccccga gcggcagcca cgatggtgaa atcgctccc 12900
 gtgaaacggt ggagctctcc ttcagcaccg tgaagcagga atacgtggtg cagaaccagc 12960
 agggcggcag cggcggcacc atcaccgcag gctacgactt caaggccaac aaagaaattt 13020
 aacggctggt tttccggcca gatgttatgt ctggctgggt ttattgtttt gattttaaag 13080
 gaatttacag tgaataaatg gcgtaacccc actgggtggt tatgtgcggt agctatgcct 13140
 tttgactgc tcctgcttcc cggatgcggc agtagcgatt cgctacttga cccctaactg 13200
 cagcggcctg gcctgagcgt gaaagcgttt tacaaggtga attctgacaa tcagaagaaa 13260
 gcggcgctca tgaagatacg tgttgagaat taatgacctt cacagaattt tttagaggtta 13320
 agcaaaatga acagacctt attcaatgaa gcgtgggttag cttttaggaa ggtgaatcat 13380

tccgtcgtcg atgtgggtag cattattggt ggaaacgttg ggaaaaatat aactgggtggt 13440
tattttcaaa atgcctgccc tattcgaatg agctatgttt tgaatgcgac aggggtccca 13500
atagcccgtg actctccgta tgcaaagggt agtgggtgccg ataataaatt ctatatttat 13560
cgcgtgaatg atatgattga ttatottact catactatgg gcaagcctga tcttattgtt 13620
aataatccga aacagagtga ctttatcggg aagaaaggaa ttatcgtagt aaaagggcat 13680
ggctggagca atgccagagg acacgttaca ttatggaatg gcagtatctg ttcagatcag 13740
tgccacttat taaatgaccc agataatgga ccatttggtc ctgaagtggg gacactgtgg 13800
atactgccgt gaaatgggta atattagtta ctttcagcat aagtggaatg ctgggtatggc 13860
agccatcctt tgcacaagaa gcattgacca cacaatatcc acagtcggaa ctcttaaaaa 13920
attgggcgct gagtcattgt ctggcattag tatacaaaga tgatgtcgtt aaaaacgatg 13980
ccagagctac ggccagtgtt taccttgaat atggtaaaaca atctgtggag atttaccatg 14040
aaattgatga gattgcgaaa aaatattcag ggttgaaata taacgggttcg atatcatcag 14100
attttaatac catgaagtgc atagatttta tccatgacag ggaattaaat gaattaatta 14160
aaaggcgtgt cgagaagtaa aattcaaaga tattaaaagc atacgttctc ttgctctgat 14220
gaatttatgg gtaagaaaga gctgtacagg aatagttaat ctgttcacct aataaagcag 14280
ataaatcagg gcttaattta ggtagttaa aggatagtag atatgtctta tgacattttt 14340
ctgaaaattg acggcattga cggcgagtca atggatgaca aacacaaaaa tgaaattgaa 14400
gtactgagct ggcgtggaa tattcatcag gaatccacca tgcacgccgg tagcggctct 14460
ggttccggta aggtctccgt cactaatctt tcatttgaac attacatcga tcgcgccagc 14520
ccgaacctgt tcaaatactg ctctccgggt aagcacattc cgcaggccat tctgggttatg 14580
cgtaaggctg gcggcaatcc gctggagtac ctcaagtaca ctttcacaga tctgattatt 14640
gcaatgggtat cgcgcagcgg aagccaggga ggggaaattg cgtctcgcga atcaattgaa 14700
ctctccttca gcaccgtgaa gcaggaatac gtggtgcaga accagcaggg tggcagcggc 14760
ggcaccatca ccgcaggcta cgacttcaag gccaaacaaag aaatttaacg gctgtttttc 14820
cggccagatt tatatctggc cggatttatt attttgattt taaaggaatt tacagtgaat 14880
gaatggcgta accccactcg gtgggttatgt gcggtagcta tgccttttgc actgctcctg 14940
ctttccggat gcggcagtag cgatgcgtta cctgacctcg aatcacagcg actcgacctg 15000
agcgtgaaag cctccgataa ggtgaatcct gacaatcaga agaaggccgc gcccatgag 15060
atacgtgttt atgaactgaa aaatgacgcc gctttcacga cagctgatta ctggtcgctc 15120
catgacaacg acaaatccgt ccttaccgac gatttagtgc gtcgcgacag ctttattttg 15180
cgtcccggcg aagagaaaaa actgcgtcgc ccgctgaatg cgcagaccac ggcaatcggc 15240

gtactggccg gataccgtaa cctggccaaa tcggtctggc gggtaacctt caaaatcccc 15300
gaagccccgg aaaaagcctg gtacagcagc ttcatatcgg ggaaaggaaa agtgcagttg 15360
gaggcggaac tggaacaaag cgccattgta attacggaac gggataaatg aattatgagc 15420
tggaatgacc gcgtagtctg gagtgaagga caattttttac tgccgcagat gtttcagcag 15480
caagagcggt atctggaaca cgtcatgcat taccgcagcc tgccgctgac cccctttttc 15540
tggggattca gccactacaa tattgatggc gaagcgctga acatcggtta actgatactg 15600
aaagaggcat cagggatttt tcctgacggc acgccgttta acgcaccgga ccacaccccg 15660
ctgccgccgc cactgaccat tctgccggag cacctgaacc agcagatttg tctggcggtta 15720
ccggtacgcg cgccgaacag cgaagaaacc acgtttgaca ataaccgga atcattggcg 15780
cgtttctcgg tacatgaaca cgacatccgc gacgccaaact cgctgggacg tggcgcgag 15840
ttattacagc tcagtcattt gcgcctgccc ctgctgccgg aaaaggcggt gacggcgccc 15900
tggattggcc tgccgttgac ccgcattacc ggggtgaacc ctgacggcg gatagatatc 15960
gaccacgacc tgatccccgc catcattaat tatcaggcca gttcactgat gtgtacctg 16020
ctgtcgtgga tcaacgatct catccggatg cgggccgatt cgctggcgga acggctgacc 16080
ggcagcgaca accacggcca tgaagcagcg gaggtctccg attacctgct gctgcaaatt 16140
ctcaatcgct ttgagccgct gctgactcac ctggcgaaaa ccccgctggc cccggagggtg 16200
ctgtaccgct acctgtccga actggccggg gaactctcca cctatgtgag tccacaaacg 16260
cgacggcccc ctgaatacaa agagtacaaa cacctgacgc cctatgccgg gttgaaatcg 16320
ctgggtgatg aggtgcagtt cctgctgaac gcggtactga tccggggcgc gcagcgcac 16380
gagctgaaag aggggactta cggcatcctg aatgcgggtg tggcccttcc cgatcttgcc 16440
gatttcagca cgctggtact ggcgataaag gcttcaatgc cgaccgatgt gctactgcaa 16500
cattttgccc ccagaccaa aatcggggcca tccgatcgcc tgccggaact gatccgctcg 16560
catctgccgg ggctggcttt gcaggttctg cctgtaccac cgcgccaaat cccgtttcag 16620
gccggataca tctattacga catccgccgc gagggagcat tgtgggaaca cattgcccg 16680
tacggcggga tggccatgca taccgccggg gaatttccgg ggctggagac agaactgtgg 16740
ggagtgcgcg ataaatgaca gacagtacc tgacgccgcc agcggcggt atgatgtcct 16800
ttttgtccac cacgccgga cataaggaca gtgaatatga aacgccggt cacaccagcc 16860
agcgcacgga actcaatgtc atcgttgaag acggtccgga cagcaaactc cggctggctg 16920
aaatcagcgc ggcggctaac ccgttgctcg ccgctgccc gcctttattg tgcgctctcg 16980
cagccatgcc cgctaaactg gatgcggccc tggtagagcc ttaccgtaat ctgctgggtac 17040

ccggttacag	gaagagtacg	accttaaaaag	ccgccagcgg	ctgtatacct	tcccgcgtga	20760
gttcgccgcc	ctcggcgagc	cgttgctgga	ggctattgaa	cagatcttcc	tcgattcaaa	20820
attcgatgcc	acgcaactga	ataacacgct	gcgcggggtg	tttttcacca	gcgccgcgca	20880
ggcgcaggcc	gacgccgtgg	ccgaccagtt	gagtatctgg	cagcgctttg	tccggggcgat	20940
aaaaaccgcc	cgtggcgaaat	cctccgcctc	tctcccacac	gctctgccgg	acggcaaccg	21000
cagctacttc	ctgcatgacc	tgctgacaca	gtttatTTTT	cgtgaagcgc	acctggtgga	21060
gccaaacctc	cagtgggcct	ggcgttaccg	cctgctgcgc	ctcggcgggc	acctgctggg	21120
actggtgctg	gcattcctgt	tgtggcaggg	gatgcagacc	agccagcaga	ccaacggcga	21180
ctatctgaat	gaaatcagcg	cccgcgcgac	ccggctggac	ggtgatgtga	aagcctacac	21240
cggtaaaccg	gcgatggctc	ccgtcccggc	actgctggac	agcgcaaggg	aactgtccgc	21300
ctggccggaa	ctggacccgg	acgcgcgcgc	gctggcctgg	cgctacggtc	tgtacagcgt	21360
accgccggta	accgacagcg	tggcgctcgt	gtacaaccgt	ctgctggatc	aactgctgct	21420
gccgccgctg	gtgaaacgga	tggagtatgt	gctggcggac	gccattgccc	gtcaggatag	21480
taaagcggcc	tacgatgccc	tgcgcatcta	tctgctactg	aatctggata	aagatcacga	21540
agataaatac	aacgcggcgg	agatccagtc	gtgggtgatt	aacgatctgg	ggaacagcga	21600
cagcgtggcc	gggttcggcg	ggcgcgccgc	cgtgctgacg	catatcgaag	cgctgtttga	21660
gggcagccgg	gtggtgcatt	caccgtatga	gaaagatgag	gcgctgatcc	gccaggcgcg	21720
ggcattcctc	gacggtcaca	ccagtaccga	gcgtatctac	gcgcggggcg	tggcggcaat	21780
ggagagcgaa	gcgccgcagg	agttcacgct	ggtacgcgcc	gtcggcgcgg	atgcgggaac	21840
ggtctttgtg	cgtagcaacg	gcgcgccgct	ggatcggggc	gtgccgggta	tttttacccg	21900
tgaaggatac	cgggagctgt	tcgacaaacg	attaccggaa	tttgtggcgg	cggcgacggc	21960
gaacgatggc	tgggtgatgg	gccgggagag	tacgccaaaa	aagctgactg	acagcctgcg	22020
cagccagata	ccggggcagg	agcagtctgt	cgcccgcgaa	gtccgccggt	tgtacctgac	22080
ggaatatgcc	cgccgctggc	aggattttct	ggacagtatc	catagtatca	acagtgccgg	22140
ggaagagggc	agttccggcc	tggcctatga	tttacaggtg	ctgcgcaccc	tggcgctcgc	22200
ggactcaccg	ctgatgcggc	tgggaaaagc	ggtggtggag	cagaccacgc	tggtgccgcc	22260
gccggacccg	caggccagac	agaaacaact	ggcgcagcgc	gcatccggta	acgcggggaa	22320
agtggtagag	acggcaaaac	tgttccagga	tattcacccg	gaagaacggc	tggaaaaaac	22380
gctggtggat	gatcgtttct	ccgcgctgcg	cgaggctcatt	gccggggcgg	cggacggcgg	22440
acagagcggc	ggtgggacga	tgcagatcgc	ttcgctgctg	accatgctca	acgagtatta	22500
caccaactg	accattgctg	atagcgcgct	ggcggcgggg	acgttgccag	cgcgcattac	22560

tgcgcgacac	aagttgcaac	tggaggcggc	gaaactgccc	gcgccgctga	aaaacatcct	22620
gctggatctg	acgaagcagg	gaacgcgcaa	aatcaacgcc	gggaccggcg	acgtgctgaa	22680
caccagatg	gaggcgatga	tgggcgacga	ctgccgtgac	gccatcgacg	ggcgctatcc	22740
gttcgccgac	agtccgcagg	aggtcagcgc	cgaggacttt	aaccgcatct	tcgccagcgg	22800
cggcgactg	gatgctttct	ggagcaaaca	actggctccg	ctggcgga	ccgccagcga	22860
cccgtggcgc	tacaaaccga	ccgaaggcaa	catgacattg	cagggggccg	atctgacgcc	22920
gtttcagcag	gcgaagcaga	tccgcagcgt	atTTTTTaa	agcgagggcg	ggaaaaaatt	22980
ttcctggctg	atgcagatta	gcgtgggtga	tatggaccgc	gccatcacgg	aactgggtgat	23040
tgatatcgac	ggccagggtg	tgcgctacgc	ccatgggtccg	gaccgtcccc	tgaaagtaac	23100
gtggccggga	ccgcgcgaacg	gctcgatggc	ggaaatcacc	gccagcccgc	gtattcgcca	23160
ggacatctca	acgctactga	ccggcgggcc	gtggggcgctg	tttcatctgc	tggacgccgg	23220
aatgggtacag	gaaacggcg	tgcgggggcg	tcaactgggtg	gaatatgatt	ttgacggctcg	23280
ccgggtgggtg	ctggaaatca	ccgccggggcg	ggatTTTTaa	ccggtcagcc	gggagctgtt	23340
gcagaacttc	agttgcccg	cgagggcgct	gtaatgcgcc	ggccagcagt	ggtatcagcc	23400
ccggcgatgt	tcggcagggt	gccggaccag	cgcgactatg	tgcgctggcg	agtgggcgct	23460
gaagagggac	gaatatggca	ggactggctg	aaccgccaga	cgtgggtggg	cagcggggcg	23520
catatcggtg	tgccgggagg	ccagacgacc	gatgaagagc	gtgacggctg	gatgcacctg	23580
tcggccgctg	gcgaacatga	tatgcccgcg	ccggaacacc	atccgctgcc	gtggagcttt	23640
gtgatgtcgc	cgggattttt	gccgatgggt	gggacgactg	actggctgac	cggagtgcctg	23700
atggcgtcac	gggatagcgt	gggacgcccc	tggccgctgg	tgatttacca	gcgctgcggt	23760
cgggagtggc	tggacgagag	tttgcaggag	acgcagggct	ggctgtactg	gctggcgagg	23820
ctggcgggcg	agcatatcac	gccggacacg	atgcggcgcg	ggcggtgac	cgagcaggtg	23880
gaccagctat	gggcgatgtg	gcagccgggg	ccgtgggtgg	cgcagtggct	gcgtgggtta	23940
cgacgtacat	cgcaacgcag	ccgggaactg	accgggctgc	cggatgaagc	accggtagta	24000
gaactgccgg	gcgtgcgtta	tctgccgtgg	ccgggctggc	cgggtaagac	gctggggcag	24060
gcaacaccgg	ggcaaggctg	gttctggcag	cagaacagtg	aagggcggtta	tgtggatgcg	24120
ctgcggttgg	tggatgaaaa	aaatagaggt	cagatgctgt	gaaaaagtac	gctgtagagg	24180
ttctgtttat	gtctgcctgt	gcaggaatgt	tcctgcccg	atttgcttgg	ggcggaaacag	24240
atgtaaacad	tgataacccg	ctggcgaggt	gtgtggatat	ccatcccgtt	catcgtcagg	24300
aaatggataa	ccttaccatt	ctgaaaacaa	cggtcacact	gaaaaaaaagc	accggtgaat	24360

—

gctggctgggtg gaaagctacc cggtagcgac ctggcaggtg cagtacgggtg aaactgattt 26280
tgattttctt cagcgactga tgcaggagtg gggcatctac tgggtggtttg agcacagcga 26340
ggacagccac acgctgggtg tggcggatgc catcagcgcc cacaaagcat gtccggactc 26400
gccgctgggtc gagtggcacc aggaagggct gaagctggac aaggagttaa tccacactat 26460
cacggcaaac gagagcctgc ggactggaca gtgggtgctg gatgatttcg attttacgaa 26520
gccacgttca ttgctggcaa acaccgtggc aaaccgcgt gaaaccgggtc atgccaccta 26580
cgagcattat gagtggccgg gagactactt cgacaagagt gaaggcgaga tgctgacgcg 26640
cattcgtatg gaagcgcagc gcagccccgg cagtcgggtg ctggggggag ggaatatccg 26700
cacactcatg accggttata ctttcacgct ggaaaactat cccaccgccc aagtcaatca 26760
ggaatatctg ctgatgcaga ctttgcgtt tgtgcaggac aacgcgcagc acagcgggca 26820
ggaccagcac ttacaccttt ccaccgctt tgaactgcac cccaccgccc aggtgttccc 26880
cccgagcgg acggtgagca aacccacac caaagggccc cagagcgcca tcgtcacccg 26940
cccggcgggc caggaaatct ggacggatca gtacgggccc gtaaagggtac agtttggttg 27000
ggatcgctac ggcaaaatgg atgaaaacag cacctgctgg atacgcgtca gctaccctg 27060
ggcgggcaaa ggcttcggga tgatccagat cccgcgtatc ggccaggaag tgctgggtgga 27120
tttcaaaaac ggcgatccgg atctgccgat catcgtgggg cgtacctaca accaggacac 27180
catgccgccc tggggactgc cgggaatggc gtcgcagagc gggatcttca gccactcgct 27240
gtatggcggg ccaacgaacg gcaacatgct gcgttttgac gacaaaacgg gcgcggagga 27300
agtgaagtcc cagcggaacg aagatctcaa caccacgggt aagaataatg aaacgcatac 27360
ggttatgggt gatcgacta aaaccattat taaaaatgaa accaacagta ttgggtgagga 27420
cagaaacacc acggtaacga agaattgacg cttttccgta aaactggcgc agacgatcaa 27480
tatcggcacc acttatcggt tagatggttg cgatcaattc acgcttcgct gcggcaatgc 27540
ggcgcttggt ttacataagg acggctccat tgagttttgt ggcaagcaac tgatgttaca 27600
taccagcgat gtcatgcaac tgattggtaa aggtattgat atgaaccggt atggcggcac 27660
agccgtaacc gccgatgata ttgccccct tctcacctct gagtgatctg aattaaacct 27720
ggagttctca tggatcgacc ataccgata caggaagggt gttttgtcct gcctgaaaca 27780
tttacggatc gcagcgtcaa ttttttatc ctggagggca atgaacgaac atcgcccagc 27840
ctgaatattt cccgcgatac gctaaaacct gatgaagacc tgcccgccta tattgaccgc 27900
cagattgcac tgatgaaaaa aaatctcggc cagcaccggg tattgtcgcg agcgctgca 27960
caggcaggaa cgggcaatga tgcccttatg ggggaacaaa ttgccgccac ccataaatcc 28020

gggaaaacgg aagtgtacca gcgtcaggcc ggggtttattg caacccttgg caagggtactg 28080
 gtcttcaccc tgaccagtcc ccgtcctttt gatgataaag cagacctact ctggaacacc 28140
 tggctggcag gctttcagcc ggataaaaac gaataatcac acggagggtgt gaccatgtat 28200
 gaagcagccc gtgtggatga tcctatctac cacaccagcg cgctcgccgg gtttcttatac 28260
 ggcgctatca tcggcatcgc cattatcgcg cttgccgcct ttgccttctt tagctgcggt 28320
 tttcttgccg ggctgattct ggggttttatg gccgatcaaa tagcctccgg ggtattgcaa 28380
 ctgggcgagg ccatcgggcg ctccatccac cacacggcag gaaaaatcct caccggttcg 28440
 gagaatgtca gcaccaacag tcgcccggcg gcgcgcgcgg tactgagtac ggtgaaatgc 28500
 gataaccata tcgcagaaaa acgcatcgcc caaggggtcgg aaaatatcta catcaacagc 28560
 cagcccgcgg cccgtaagga tgaccacacc gaatgcgacg cggtgattga agacggttcg 28620
 ccgaatgtgt ttctcggcgg cggcacacag acgggtactgg aaatcagttc tgaaattccg 28680
 gactggctgc gcaagggtgg ggatgtattg tttgtcgtgg cgagtctgct cggcgggctg 28740
 gccggggcgt ggcggcaggc ggcaaagctg gggacgaaat ttggcactaa atgtgccgct 28800
 aagtttatcg gcggggagct tgtcgggatg gccgtgggtg aggctatcag cgggctgttc 28860
 agcaatccgg tggatgtgac caccgggcag aaaatcctgc tgccggaaac ggacttcacc 28920
 ctgcccggtc gcctgccggt cacctgctcg cgtttttacg ccagccacct ggaaactgtg 28980
 ggactgttgg gacggggctg gcggctgaac tgggaaacca gcctgcgcga tgacgatgaa 29040
 cacatcacgc tgaccggcgt acagggggcg gaactgcgtt acccgaaaac gatgctgacg 29100
 cccggccacc agatatattga cccggaagaa cagttatacc tcagccgcct gcatgacggg 29160
 cgttacgtgc tgcattacac cgatcgacgc cattacgtat ttggtgattt tgacagtacg 29220
 ggcattggcat acctgctgtt tatggagacg ccgcaccgcc agcgcatgtt cttcggggcac 29280
 gaaggaggca gactggtacg gatagcctcc agcagcgggc atcacctgtt actgcaccgc 29340
 acacagaccc cggcagggga gcggctgtcg cgaattgaac tgggtgcaggg cggcaccctg 29400
 ggcaatctgg tggagtaccg gtatgacgat aacgggtcaac tgaccggcgt ggtgaaccgg 29460
 gcgggaacgc aggtgcgtca gtttgcttat gaaaacgggc tgatgacggc gcacagcaat 29520
 gcgacggggg tcacctgccg ctaccgctgg caggaaactcg acggcgcgcc gcgcgtgacg 29580
 gagcacgaca ccagtgcagg cgaacattac cgctttgact atgattttgc cgcaggcacc 29640
 accaccgtca ccggcaggca gggggagaca tggcagtggt ggtacgacag ggaaacgtat 29700
 atcaccgcgc accggacgcc gggcgggtgga atgtaccgct tcacgtacaa cgaagaccac 29760
 ttcctgtca acattgagct gcccgcggt cgcacgggtg cgtatgaata tgacatccag 29820
 aaccgggtgg tgaagacgac agatccggaa ggccgggtga gcgagacgca gtggaaccgg 29880

gagttcgacg aaatcacgcg cacggcgctg gacgatgacg ctgtctggaa aacgcagtac 29940
aacgcccacg gccagccagt gcaggagacg gacccggaag ggcgggtgac gcagtacgct 30000
tacgatgaac aggggacagat gtgcagccgg acggatgcgg cgggaggcac gcaggggcggc 30060
gtgcggcggg agacgcagca gcgggatgcg ctggggccgtc tgttacggac ggagaatgaa 30120
cacggccagc ggacgttcag ctacaaccgg ctggaccaga taacggcagt gacgctcacg 30180
cccacggagg cggggcaaca gcagcaccgg atgcaggccg acacggtgcg ttttgagtat 30240
gaccgcagcg gctggctgac ggcggagcac gcggggaacg gtagcatatg ttatcagcgc 30300
gacgcgctgg gcaaccgcag ggacatcacg ctgccggacg ggcagcacct gacgcattcg 30360
tattacggga gcgggcatct gttacagacg gcgctggacg gcctgacggg gagcgagtat 30420
gagcgcgaca gcctgcaccg tcagataatg cgcacgcagg ggcagcttgc gacgtacagc 30480
ggctatgacg acgacgggct gctgagctgg cagcgcagtc tggcgctccg cagtgcacct 30540
gttcttccctg gccagcgccc ggcgcggcag ggctgcgtga cgtcgaggga ctattactgg 30600
aacaaccacg gcgaggtggg cacgattgac gacggcctgc gtggcagcgt ggtgtacagc 30660
tatgcagaaa gcgggttacct gaccggggcg tcaggtcaga tgtatgacca tgaccgttat 30720
tattacgata aggcggggcaa cctgctggat aacgaagggc agggagcggg gatgagcaac 30780
cggctgccgg gctgtggtcg tgaccgttac ggctataacg agtggggcga gctgaccacg 30840
cggcgcgacc agcaactgga gtggaacgcg caggggcagc tgacgcgggt catcagcggc 30900
aacacggaga cgcactacgg ctacgatgcg ctggggaggc gaaccgcga ggcgacgtac 30960
ggscggcaca cggggccatac ggcgcggagc cggacggact ttgtgtggga ggggttcagg 31020
ctgttgacgg agaactgca gcagcagggc tggcggacct atctgtacga tgcggaacag 31080
ccgtacacgc cgggtggcgag cgtgacgggg cggggagaaa gcaggcaggt gtggtattac 31140
cacacggatg tgacgggcac gccgcaggag gtgacggcgg cggacggaac gctsgtgtgg 31200
gcgggggtata tcaggggggt tggagagaat gcggcggaca tcagcaacag cggggcgtag 31260
ttccaccagc cgctgcggct gccggggcag ttttttgacg acgagacagg gctgcattac 31320
aatctgttca gatattatgc accggagtgt ggacggtttg tcagtcagga tccgatcggg 31380
ctgagggggc gggttaaacct ttatcagtat gcgccaaatc ctctcaaata tatagacca 31440
cttggtttaa ccgcgactgt tgggcgatgg atggggcctg cggaatatca gcaaagtctt 31500
gatactggga cagtagtaca aagttcaaca gggacaactc atgttgcta cctgctgat 31560
atagatgctt ttggtaagca agcaaaaaat ggtgctatgt atgttgaatt tgatgtgcct 31620
gaaaaatcat tagtacctac aaatgaagga tgggcaaaaa tagtagggcc agattctatc 31680

gaagggcgat tagctaaacg caaagggttg cctgttcctg aaatgccaac agcagaaaac 31740
ataactgtaa ggggcgagaa aattaatggg gaagttgaag caaaatgcta aataaattta 31800
aattgtgggt gagcaaacat actgattata cggtaattca taatgaaaat gatttatctt 31860
acagtattat tatagatttt gaagatgacc ggtatatatc aagatttact gtatgggatg 31920
acctaagctg tatgtcagaa gtaatggatg tggatactgg tttatataaa ttaaacaaga 31980
gaaacgaatt ttctacattt gatgaacttc tggatatatt tgatgatttt atgataagta 32040
ttaaataata gttggccggg taagaagtta actcttcccg gctgttttat tatctaacc 32100
ccatcaatcc ggagacgcgc taccggtacg atgcgctggg caggcgggtg agcaaggcga 32160
cgtacgggcg gcacacgggc catacggcgc ggagccggac ggactttgtg tgggaggggt 32220
tcaggctgtt gcaggagaac gtgcagcagc agggctggcg gacatatctg tacgatgcgg 32280
aacagccgta cacgccggtg gcgagcgtga cgggaaaggg agaaagcagg cagggtgtgg 32340
attaccacac ggacgtgacg ggcacgccgc aggaggtgac ggcggcggac ggaacgctgg 32400
tgtgggcggg gtatatcagg gggtttgag aaaatgcggc ggacatcagc aacagcgggg 32460
cgtactttca ccagccgctg cggctgccgg ggcagtattt tgacgacgag acagggctgc 32520
attacaatct gttcagatat tatgcaccgg agtgtggacg gttcgtcagt caggatccga 32580
ttgggctggc ggggggggct gaatctttac cagtatgcgc ctaatccgat tagatggatc 32640
gaccttttag gacttgctat cctggagcat caatctaatt ttgatgcggc aaggagaacc 32700
ggatttgaaa atgcgggtat gacaaaccct gaggatgtca ctttctcgaa agtcgatccc 32760
aaaactggta ctgttggtga gtttaaaggt ccaaagggg ctaaagttgc ttatgatgca 32820
cctcatgcag atatggatgt gacagcaggc catgataaac cacatgttgg ttggcaatcc 32880
gcaggaaaaa gaggttccgg aggagctaata agaggtaata ttacttatga tggcccacaa 32940
catccgcac gctctgactc taaggagat gataaatgtt aaattcaaat atgtctgaac 33000
ttagaatcga actggagaat gcgattaaaa atctcggtat tcatgattat cgtgtcgata 33060
aaccgaaca aatcgtttct gagataaaag agatatatgt taatggtaat cctagaacct 33120
gggtggttatc attaaaacat agacaatatg tcttttctta taccgataat tctggatata 33180
aaaacatatc acaaatagta agtaaacaac tcaatgaaag caatgtaatc acaaacata 33240
tatttttgat tgctgatgaa gataatgagc aaatatatgt atataacgtt cctcttaact 33300
ccctgcctga aattatagaa aattgcagat attttgaata ttatgttgca gatcatgaac 33360
tatcttggct tatatgtgaa aatgatcatg gtgatttgat tgtatgctca accattaagt 33420
aaagcgcgag tgctctttag cgatatagtt gcccatatct aggcgttact agccgaagat 33480
ggcgcgattg tctggcaggg gaaacagcaa ttctgaggtc aggaagatag cataaccat 33540

cggtgtgatg aggatgaatt taatgtatat cgagaaacgg ttgggttttat catgggtgaa 35400
atgcttatta aaataatgaa tccattatat gaaaaacatc cagaaataaa accaaaagga 35460
ttgaaacaaa acatctggaa ccggatgaat aatgtgtaaa agccggaggg gttatctttt 35520
ccccgctttt tattatcaat tactcattaa ctctgttcc gttcttttgc gtttaatcac 35580
cggaatatct ccggtattgt tcagcgcccc ggaaatgttt ttaaccactg ttctgcactc 35640
cgtttattaa tgcgggttac gcccatccct tcaatacagc caaagagtcc gtgggtatgc 35700
tgcggcgtga tcacgatgca atccctcatt acccgcacct tgacgggcat cccgttaata 35760
aaccgcgctt gcggcatcca ttcaccgtac aaacgaactg aaagtctctc aacgcgtgag 35820
tatgtaagta tcccgcataa tcgagccatt cacatttaga gatcatccga cataatcaat 35880
ctgccaacgc aggagatcgc tatgcgtaaa gcccgattta ctgcgcacca gatcatcgct 35940
gtgattagat cagtcgaatc cggacggact gttaaagatg tctaccggga gcccggtatt 36000
tctgaagcca ccagggacaa ctggaggtct ggatacggcg gcagggatac gcgtggaatc 36060
acaaaaggct ccaccgtatt tactgtctgc tcaagctgaa ttttcgccgt aagggtaaac 36120
aacggctgcc ggtacgcaat ccctcgccac tggtcacgcc ggaagcgctg aaccagagct 36180
ggctctgtgg cgctgttttc gcacgttcaa tgttgttgat gactgtaatc gtgaagcgtt 36240
gtcgattgaa atcgatctga atctgccagc tctgcgagtg gtccgtgtac tcgacaggat 36300
tacagcaacc gcggttatct ggccatgctg cgtatggata agggaccgga atttatctcg 36360
ctggcactgg ctgaatgggc aaagaaacat gcagttaaagc tggcgtttat ccagccgggt 36420
aagccgaaga aaaacgtttt catcacgcg ctttaaccgga cataccgtac agaaatactc 36480
aattcttatt tgttcagaac gctgaatgag gtgtgggaaa ttacggataa agggttatca 36540
gaatataact gcgaacgtcc acatgaatcg cggaacaata tgataccgaa ggaataccgc 36600
caataacggt atctggccgg aatcttaaaa atgcatggaa ctaaaacggg tctatttaca 36660
ggggcacctg cgatgaattt cgctgcactg aaaagcgata ccggatgaga gctgcttcaa 36720
attaatgtgc catgttcacg gggaggttgt gcgacgtttg cataatccag caagaactga 36780
aaggaagggg agagcttttt catgcctgta taatcagtct ggctgtgtc agtcagctct 36840
tagtgttgag actctcgttg gagcgttata attgcttttc tgtttcggaa aacaagattt 36900
tccattaaag atcttccctg cgaggaaaag ttaactaata atcttaccgt cgagtttagga 36960
gatgtatggt taaatataaa caatgttgca acgatgcctg ataattatcc tctcttcgaa 37020
gataagtttc ccacaccag tgtagtaggt gtcatggtaa tgttatcact tgaatgtaaa 37080
tggaaggat aattgctttt tgactggcat tctattccac cctgacaaca cgatgttaac 37140
atcaaacctg tttatattgg caataacgca atttttttca gattaagagg tgctctgata 37200

Substitute sheet (Rule 26)

tca gta tat gcc gtc aac cag caa tta aat tca gcc act aaa tta ttc	38059
Ser Val Tyr Ala Val Asn Gln Gln Leu Asn Ser Ala Thr Lys Leu Phe	
190 195 200	
agc gtg aag ctg ggg gct aca cga gtg att tat cac gct ggt acg gct	38107
Ser Val Lys Leu Gly Ala Thr Arg Val Ile Tyr His Ala Gly Thr Ala	
205 210 215	
gga gcg acg ctc tcg gtg agc aac ccg cag aat tac cct att ttg gtt	38155
Gly Ala Thr Leu Ser Val Ser Asn Pro Gln Asn Tyr Pro Ile Leu Val	
220 225 230 235	
cag tct tca gtc aaa gca gca gac aaa agt tcg cct gct ccc ttt ttg	38203
Gln Ser Ser Val Lys Ala Ala Asp Lys Ser Ser Pro Ala Pro Phe Leu	
240 245 250	
gtg atg ccg cct cta ttt cgt tta gaa gca aac cag cag agt caa ctg	38251
Val Met Pro Leu Phe Arg Leu Glu Ala Asn Gln Gln Ser Gln Leu	
255 260 265	
cgt att gtc cgt act ggt ggt gac atg cca acg gat cgt gag act tta	38299
Arg Ile Val Arg Thr Gly Gly Asp Met Pro Thr Asp Arg Glu Thr Leu	
270 275 280	
cag tgg gtc tgt ata aag gcg gta cca ccc gaa aat gaa ccg tcg gat	38347
Gln Trp Val Cys Ile Lys Ala Val Pro Pro Glu Asn Glu Pro Ser Asp	
285 290 295	
aca cag gct aag ggc gcg acc ctt gac ctc aat ttg tcc atc aac gcc	38395
Thr Gln Ala Lys Gly Ala Thr Leu Asp Leu Asn Leu Ser Ile Asn Ala	
300 305 310 315	
tgt gat aag ctg att ttc cgc ccg gat gcc gtg aag ggg acg ccg gaa	38443
Cys Asp Lys Leu Ile Phe Arg Pro Asp Ala Val Lys Gly Thr Pro Glu	
320 325 330	
gat gtt gca gga aat tta aga tgg gtg gag acg ggc aac aaa ctt aag	38491
Asp Val Ala Gly Asn Leu Arg Trp Val Glu Thr Gly Asn Lys Leu Lys	
335 340 345	
gtg gag aac ccc acc ccg ttt tac atg aat tta gcc tct gtc aca gta	38539
Val Glu Asn Pro Thr Pro Phe Tyr Met Asn Leu Ala Ser Val Thr Val	
350 355 360	
ggg gga aag ccc att aca ggg ctt gag tat gtc ccc ccc ttt gct gac	38587
Gly Gly Lys Pro Ile Thr Gly Leu Glu Tyr Val Pro Pro Phe Ala Asp	
365 370 375	
aaa aca cta aat atg cca ggt agt gcc cat ggt gat atc gag tgg aga	38635
Lys Thr Leu Asn Met Pro Gly Ser Ala His Gly Asp Ile Glu Trp Arg	
380 385 390 395	
gtt att aca gac ttt ggt ggt gaa agt cat ccg ttc cac tac gtt ctt	38683
Val Ile Thr Asp Phe Gly Gly Glu Ser His Pro Phe His Tyr Val Leu	
400 405 410	
aaa taa atccaggggc ttagcggcag aaa atg aag ttc aaa caa cct gcc ttg	38736
Lys Met Lys Phe Lys Gln Pro Ala Leu	
415 420	
cta ctg ttc atc gcg gga gtg gtt cat tgc gca aat gcg cac act tac	38784

Leu Leu Phe	Ile Ala Gly Val Val	His Cys Ala Asn Ala His Thr Tyr	
	425	430	435
aca ttc gat gca tca atg ttg ggc gat gca gcg aaa ggg gtt gat atg			38832
Thr Phe Asp Ala Ser Met Leu Gly Asp Ala Ala Lys Gly Val Asp Met	440	445	450
tcg ctc ttt aac cag ggg tta caa cag cca ggg act tat cgc gtg gac			38880
Ser Leu Phe Asn Gln Gly Leu Gln Gln Pro Gly Thr Tyr Arg Val Asp	455	460	465
gtg atg gtg aat ggt aaa cgt gtc gac acc cgt gat gtg gtg ttc aaa			38928
Val Met Val Asn Gly Lys Arg Val Asp Thr Arg Asp Val Val Phe Lys	470	475	480
			485
ttg gaa aag gat ggg caa gga acg cct gtt ctg gct cct tgt ttg acg			38976
Leu Glu Lys Asp Gly Gln Gly Thr Pro Val Leu Ala Pro Cys Leu Thr	490	495	500
gtc agt cag ctt tca cgc tac ggc gta aaa acg gaa gat tac cct cag			39024
Val Ser Gln Leu Ser Arg Tyr Gly Val Lys Thr Glu Asp Tyr Pro Gln	505	510	515
ttg tgg aaa gca gca aag ccc cca gat gag tgt gcg gat ctg acc gcc			39072
Leu Trp Lys Ala Ala Lys Pro Pro Asp Glu Cys Ala Asp Leu Thr Ala	520	525	530
att cca cag gct aaa gcg gta ctg gat atc aat aat cag caa ctg caa			39120
Ile Pro Gln Ala Lys Ala Val Leu Asp Ile Asn Asn Gln Gln Leu Gln	535	540	545
ctg agt att ccg cag ttg gcg ttg cgt ccg gaa ttt aag ggg atc gct			39168
Leu Ser Ile Pro Gln Leu Ala Leu Arg Pro Glu Phe Lys Gly Ile Ala	550	555	560
			565
cca gaa gat ctt tgg gat gat ggt att ccg gcg ttt ctg atg aac tac			39216
Pro Glu Asp Leu Trp Asp Asp Gly Ile Pro Ala Phe Leu Met Asn Tyr	570	575	580
agt gcg agg aca acg cag acg gat tac aaa atg gat atg gtg ggg cgt			39264
Ser Ala Arg Thr Thr Gln Thr Asp Tyr Lys Met Asp Met Val Gly Arg	585	590	595
gac aac tct tcc tgg gta caa ctg caa ccg gga atc aat ata ggt gcg			39312
Asp Asn Ser Ser Trp Val Gln Leu Gln Pro Gly Ile Asn Ile Gly Ala	600	605	610
tgg cgt gtc cgc aat gcg acc agc tgg cag ccg agt agt caa ctg tcg			39360
Trp Arg Val Arg Asn Ala Thr Ser Trp Gln Arg Ser Ser Gln Leu Ser	615	620	625
ggg aag tgg cag gca gca tat acc tat gct gag cgt gga ctg tac tca			39408
Gly Lys Trp Gln Ala Ala Tyr Thr Tyr Ala Glu Arg Gly Leu Tyr Ser	630	635	640
			645
cta aaa agt cgt ctg act ctg ggg caa aag act tcg cag ggg gag ata			39456
Leu Lys Ser Arg Leu Thr Leu Gly Gln Lys Thr Ser Gln Gly Glu Ile	650	655	660
ttt gat agt gtg cca ttt acc ggt gtg atg ttg gca tcg gat gac aac			39504
Phe Asp Ser Val Pro Phe Thr Gly Val Met Leu Ala Ser Asp Asp Asn			

WO 00/73336

PCT/SE00/01079

25

665					670					675						
atg	gtg	ccc	tac	agt	gag	cgg	cag	ttt	gct	ccg	gta	gtg	cgt	ggg	att	39552
Met	Val	Pro	Tyr	Ser	Glu	Arg	Gln	Phe	Ala	Pro	Val	Val	Arg	Gly	Ile	
		680					685					690				
gcc	cgc	acg	cag	gct	cgg	gtg	gag	gtc	aaa	cag	aat	ggg	tac	acc	att	39600
Ala	Arg	Thr	Gln	Ala	Arg	Val	Glu	Val	Lys	Gln	Asn	Gly	Tyr	Thr	Ile	
		695				700					705					
tac	aac	acc	act	gtg	gcg	ccc	gga	ccg	ttt	gca	ctg	cgg	gat	ctg	tcg	39648
Tyr	Asn	Thr	Thr	Val	Ala	Pro	Gly	Pro	Phe	Ala	Leu	Arg	Asp	Leu	Ser	
710					715					720					725	
gta	aca	gac	agt	agt	ggg	gat	ctg	cat	gtc	acc	gtg	tgg	gag	gcc	gat	39696
Val	Thr	Asp	Ser	Ser	Gly	Asp	Leu	His	Val	Thr	Val	Trp	Glu	Ala	Asp	
				730					735					740		
ggc	agt	aca	caa	atg	ttt	gtg	gtg	ccg	tat	cag	acc	ccg	gcg	ata	gca	39744
Gly	Ser	Thr	Gln	Met	Phe	Val	Val	Pro	Tyr	Gln	Thr	Pro	Ala	Ile	Ala	
			745					750					755			
ctg	cac	cag	gga	tat	ttg	aag	tac	agc	ctg	ttg	gcg	ggc	cga	tac	cga	39792
Leu	His	Gln	Gly	Tyr	Leu	Lys	Tyr	Ser	Leu	Leu	Ala	Gly	Arg	Tyr	Arg	
		760					765					770				
tcg	tca	gac	tct	gca	acg	gat	aag	cgg	cag	atc	gcg	cag	gct	acg	ttg	39840
Ser	Ser	Asp	Ser	Ala	Thr	Asp	Lys	Arg	Gln	Ile	Ala	Gln	Ala	Thr	Leu	
	775					780					785					
atg	tat	ggg	ctg	ccg	tgg	aat	ctc	act	gca	tac	ggc	ggg	ata	cag	agt	39888
Met	Tyr	Gly	Leu	Pro	Trp	Asn	Leu	Thr	Ala	Tyr	Gly	Gly	Ile	Gln	Ser	
790					795					800					805	
gca	acg	cat	aat	caa	gct	gca	ttg	ctt	ggg	ttg	ggg	gga	tct	ctc	ggg	39936
Ala	Thr	His	Asn	Gln	Ala	Ala	Leu	Leu	Gly	Leu	Gly	Gly	Ser	Leu	Gly	
				810					815					820		
cgg	tgg	ggg	agt	tta	tct	gtc	gat	gga	agc	gac	aca	cac	agt	cag	cgt	39984
Arg	Trp	Gly	Ser	Leu	Ser	Val	Asp	Gly	Ser	Asp	Thr	His	Ser	Gln	Arg	
			825					830					835			
cag	ggg	gag	gcg	gta	cag	caa	gga	gcc	tcc	tgg	cga	ctg	cgt	tac	agc	40032
Gln	Gly	Glu	Ala	Val	Gln	Gln	Gly	Ala	Ser	Trp	Arg	Leu	Arg	Tyr	Ser	
		840					845					850				
aac	cag	ctg	act	gcg	acg	ggg	aca	aat	ttt	ttt	ctg	acg	aga	tgg	cag	40080
Asn	Gln	Leu	Thr	Ala	Thr	Gly	Thr	Asn	Phe	Phe	Leu	Thr	Arg	Trp	Gln	
		855				860					865					
tat	gcc	tcg	cag	ggc	tat	aac	acc	cta	tcc	gat	gtg	ctc	gac	agt	tat	40128
Tyr	Ala	Ser	Gln	Gly	Tyr	Asn	Thr	Leu	Ser	Asp	Val	Leu	Asp	Ser	Tyr	
870					875					880					885	
cga	cat	aat	ggc	aac	cgt	cta	tgg	tcg	tgg	cgg	gaa	aat	ttg	cag	ccg	40176
Arg	His	Asn	Gly	Asn	Arg	Leu	Trp	Ser	Trp	Arg	Glu	Asn	Leu	Gln	Pro	
				890					895				900			
agc	tcg	cgt	act	acc	ctg	atg	ttg	agt	cag	tca	tgg	ggg	agg	cat	ttg	40224
Ser	Ser	Arg	Thr	Thr	Leu	Met	Leu	Ser	Gln	Ser	Trp	Gly	Arg	His	Leu	
			905					910					915			

ggc aat ctg agt tta acc ggt tcc cgt acc gac tgg cgt aat cgc ccc	40272
Gly Asn Leu Ser Leu Thr Gly Ser Arg Thr Asp Trp Arg Asn Arg Pro	
920 925 930	
ggg cat gat gac agc tac gga ctg agt tgg gga acc tct atc gga ggg	40320
Gly His Asp Asp Ser Tyr Gly Leu Ser Trp Gly Thr Ser Ile Gly Gly	
935 940 945	
ggc tcg ctg tca ttg aac tgg aat caa aac aga acg ctg tgg cgc aat	40368
Gly Ser Leu Ser Leu Asn Trp Asn Gln Asn Arg Thr Leu Trp Arg Asn	
950 955 960 965	
ggc gcg cac cgt aaa gag aac ata acc agc ctg tgg ttc agt atg cca	40416
Gly Ala His Arg Lys Glu Asn Ile Thr Ser Leu Trp Phe Ser Met Pro	
970 975 980	
tta agc cgc tgg acg ggg aat aat gta agt gct agt tgg cag atg act	40464
Leu Ser Arg Trp Thr Gly Asn Asn Val Ser Ala Ser Trp Gln Met Thr	
985 990 995	
tca cca tca cac ggt ggt cag acg caa caa gtg ggg gtc aac gga gag	40512
Ser Pro Ser His Gly Gly Gln Thr Gln Gln Val Gly Val Asn Gly Glu	
1000 1005 1010	
gca ttc agt cag caa ctg gat tgg gag gtg cgt cag agt tac cgt gcc	40560
Ala Phe Ser Gln Gln Leu Asp Trp Glu Val Arg Gln Ser Tyr Arg Ala	
1015 1020 1025	
gat gcc ccg cca ggt ggt ggt aat aac agc gca ttg cac ttg gca tgg	40608
Asp Ala Pro Pro Gly Gly Gly Asn Asn Ser Ala Leu His Leu Ala Trp	
1030 1035 1040 1045	
aat ggg gat tac ggc ctg tta ggt ggt gac tat agc tac agc cgg gcg	40656
Asn Gly Asp Tyr Gly Leu Leu Gly Gly Asp Tyr Ser Tyr Ser Arg Ala	
1050 1055 1060	
atg cgc cag atg gga gtc aat atc gcg gga ggt ata gtt atc cac cat	40704
Met Arg Gln Met Gly Val Asn Ile Ala Gly Gly Ile Val Ile His His	
1065 1070 1075	
cat ggt gtg acg ctg ggg caa cct ttg caa ggc tca gtg gcg ctg gtt	40752
His Gly Val Thr Leu Gly Gln Pro Leu Gln Gly Ser Val Ala Leu Val	
1080 1085 1090	
gaa gcg cca ggg gcc tcg ggg gtg cca gtt ggc ggc tgg cct ggc gtt	40800
Glu Ala Pro Gly Ala Ser Gly Val Pro Val Gly Gly Trp Pro Gly Val	
1095 1100 1105	
aag acg gat ttt cgt ggc gac acc aca gtg ggc aac ctg aac gtc tat	40848
Lys Thr Asp Phe Arg Gly Asp Thr Thr Val Gly Asn Leu Asn Val Tyr	
1110 1115 1120 1125	
cag gag aat aca gtc agc ctc gat ccg tcg cga cta ccg gat gac gca	40896
Gln Glu Asn Thr Val Ser Leu Asp Pro Ser Arg Leu Pro Asp Asp Ala	
1130 1135 1140	
gag gtc aca caa acc gat gtg cgc gtg gtg cca acc gaa ggg gcg gtg	40944
Glu Val Thr Gln Thr Asp Val Arg Val Val Pro Thr Glu Gly Ala Val	
1145 1150 1155	

WO 00/73336

PCT/SE00/01079

27

gtg gaa gcg aag ttt cac act cgc atc ggg gcc agg gca ctg atg acg	40992
Val Glu Ala Lys Phe His Thr Arg Ile Gly Ala Arg Ala Leu Met Thr	
1160 1165 1170	
ctg aaa cgg gaa gat ggt agc gcc att cct ttc ggg gcg cag gtt aca	41040
Leu Lys Arg Glu Asp Gly Ser Ala Ile Pro Phe Gly Ala Gln Val Thr	
1175 1180 1185	
gtc aat ggg cag gat ggc agt gct gct ctg gtg gat act gat agc cag	41088
Val Asn Gly Gln Asp Gly Ser Ala Ala Leu Val Asp Thr Asp Ser Gln	
1190 1195 1200 1205	
gtt tat ctc act ggt ttg gcg gat aag ggc gaa ctg acg gtg aaa tgg	41136
Val Tyr Leu Thr Gly Leu Ala Asp Lys Gly Glu Leu Thr Val Lys Trp	
1210 1215 1220	
gga gca cag caa tgt cgg gtt aac tac cgc cta cct gcc cac aag gga	41184
Gly Ala Gln Gln Cys Arg Val Asn Tyr Arg Leu Pro Ala His Lys Gly	
1225 1230 1235	
atc gcg ggc ttg tat caa atg agc ggt ctc tgc aga tag ccgattctga	41233
Ile Ala Gly Leu Tyr Gln Met Ser Gly Leu Cys Arg	
1240 1245 1250	
aggagagaat a atg tgg atg aaa ata cag cga gtg aaa acg gtt atc tat	41283
Met Trp Met Lys Ile Gln Arg Val Lys Thr Val Ile Tyr	
1255 1260	
agc gta agc tta ctg gtc gct gcc agt agc ttg gtg ccg ata gcg aac	41331
Ser Val Ser Leu Leu Val Ala Ala Ser Ser Leu Val Pro Ile Ala Asn	
1265 1270 1275	
gcc gca gaa aaa ctt cag aca acg cta cgt gta ggt act tac ttt cgg	41379
Ala Ala Glu Lys Leu Gln Thr Thr Leu Arg Val Gly Thr Tyr Phe Arg	
1280 1285 1290 1295	
gct ggg cac gtg cca gat ggg atg gtg ctt gcg caa gcc tgg gtg act	41427
Ala Gly His Val Pro Asp Gly Met Val Leu Ala Gln Gly Trp Val Thr	
1300 1305 1310	
tat cac ggc agt cac agc ggg ttt cgg gta tgg agc gat gag caa aag	41475
Tyr His Gly Ser His Ser Gly Phe Arg Val Trp Ser Asp Glu Gln Lys	
1315 1320 1325	
gcg ggt aac acg cct acc gta ttg ctg ctg agc ggg caa cag gat cct	41523
Ala Gly Asn Thr Pro Thr Val Leu Leu Ser Gly Gln Gln Asp Pro	
1330 1335 1340	
cgc cat cac att cag gtt cgc ctg gag ggc gag ggg tgg caa cca gat	41571
Arg His His Ile Gln Val Arg Leu Glu Gly Glu Gly Trp Gln Pro Asp	
1345 1350 1355	
acg gtg agt ggt cgt ggc gcc att tta aga acc gct gca gat aac gcc	41619
Thr Val Ser Gly Arg Gly Ala Ile Leu Arg Thr Ala Ala Asp Asn Ala	
1360 1365 1370 1375	
agt ttc agt gtg gtc gtt gat ggc aat cag gaa gtg cct gcg gac acc	41667
Ser Phe Ser Val Val Val Asp Gly Asn Gln Glu Val Pro Ala Asp Thr	
1380 1385 1390	
tgg acg ctg gat ttt aag gcc tgt gca ttg gcg cag gag gat acg tag	41715

Trp Thr Leu Asp Phe Lys Ala Cys Ala Leu Ala Gln Glu Asp Thr
 1395 1400 1405

ccgtctgttc cactcatact tcctgtatca gtgaataagg cgaataactgt atatacagggt 41775
 cgaggatgta ttgaaaatct actcaaggca aatatgggtta gcaatcttca ttgcagaaaa 41835
 atcaattgct gattgatttt aaaatctgaa attcatcttt tttgtaggga gaggaatatt 41895
 atgtttcgta aaataattta acggacactg atgcctttat tatttctgtt ttgtggattt 41955
 gtttgattca ctcttttgtg agcggggcct tataagcccc tgggtataagg tttgtttatc 42015
 tgtcagggtat tcatatgtga tatttaagat tttatgttta gggggctttt tcttgtcccc 42075
 tcaggcgtaa aaataattta ttttttacat aaaggaataa agcatatgtc ttatgcacga 42135
 catttaccgg tattaatgta tcaccatgtc agtgataaac ccggacagat aaccttatct 42195
 ccccgtagct tccgggcgca gatgaaatgg ctggccgaat ctggctggaa aaccgttact 42255
 gcctgcagag gtggaagctt tttatcatgg tgcaagattg cctcgtaaaa gcgtcatgct 42315
 gacctttgat ggcggctggc tggataactg gttgcagggt tttccgggtc tgcaggagct 42375
 taatctgcat gcgcatctct ttcttgtgac cagtttgatc agtgacggac cgggccgtat 42435
 tcctgcaggc gaaccgggtg actctcatga tgagtgtcaa atgctgggtta aacaaggccg 42495
 ggctgatgag gtcattgctg gctggtcaga ggtccgggag atgcacctca gtggccttgt 42555
 tgagtttcac tcgcacacgc acaccaccg acgctgggac cagaagcctg tgtcccgtaa 42615
 tccgtcggat ttgcttcgtg tcgatattct tcttagtctg aagcggatga gggagatgct 42675
 gggttattgc agtcagcatc tgtgctggcc tgagggtctg tattgttctg actatattca 42735
 tgtggctgaa gagttggggt tcacatacct gtataccaca gaaaggcgta tgaacaatcc 42795
 agtcattcgg tcacagcgta ttggtcgtat caacgcaaag gagcgaaaga atgtgggctg 42855
 gctgaaacgt cgtctgtttt atcacaccac gcccgattt tcttcgctgc tggcccggca 42915
 taagggggca cgtcggatag ctgactgagc cagagaccag gatgaaagtg ctgcataccg 42975
 aatcctcccc catcatcggc gggcaggggt tgtaggctat atcccaaag atgggtgctga 43035
 ttttgagcag aaataagata acaccggagg ccgtgataga cccggttgcg cagaatctgg 43095
 aggcagcagg tcaactggcg agggcttcta cccgctggct tttggttatg ggagattttg 43155
 aatgtacaga ggcccagcgg gagtggttgt tgttgcgccg gaattattgc cttgcgcaga 43215
 tatcttcccc tggcgccagt aaagctggat atcagcgacg tgacgaaatg agaggagcct 43275
 gcacataggt ttgtgttttg agtaagtgtc ttatttaaac cgtctgttct gtttcctccg 43335
 ctttcacaaa taatgtcgag ccgggtgggg gactcaagta agaataatct ggcgatgttt 43395
 tgcttgtttc cacgggatac tttgttaggt gaacgataca attaatgcgc tccagactgc 43455

gtcacagaaa aggagaatgc ctttgcagaa cgggcgcatg acaatgaacc actttattat 43515
tggatatctgg ttcggcgaca accttgacga tcacgtctga gatagggtat ttacactgag 43575
tggtaaacag gttcaattag taaccggaga tggatgcaaa atcatgatcg attcagatgc 43635
tatttttgcag ccaatagatt ttttattaag atgatatcaa ggacattgag gcacacaatg 43695
acgtatcagt aagtcgttga tagctcattt gatataagaa tttcttttat caacggaaga 43755
taatgatgga actgatcaat aatcgtggta tgcgagactg gatgatattt attaaagtgg 43815
cggaagtagg gaatctttcc cgggctgcgc gggaattaga tattagcatt tctgctgtca 43875
gtaaatcgct tagtcgcctt gagaattcta ttgagggttac tttacttcgg cgcgattcac 43935
atcacttaga actgactgga gctggtcaga cagcctatgc aagcatgaaa aggataacat 43995
cttcctttca gtccttgctg gatgaattgc gaaatccgga taaaattatc agagggagta 44055
taaaattttc ggctccggct attgtctgtg agtttcttgc caataagtgg atatgggaat 44115
ttacagctag ctatccggat acaaaaatct acctggattc acgagagcgt agcgattttt 44175
ttagtaaatc cctggagttt gatgagctgg tttttaaaag tggcataatc gaaagtgagg 44235
atctcgtgta tcgaaagata agccctttaa agttggttct ttgtgagat cggaaatata 44295
tcagaaaata tggcaggatc tcacaccctg gcgatttgga aaatcacatt attgtgggtc 44355
ttcacaacca tggctcttcc ggacctctta ctcttttccg tcaggatgaa tcatacacta 44415
ttagtgggcg tgtaaatgtt catttatctt ccaataatct tttgagtgtt ctttaatttg 44475
ttttagaagg aaaggggtatc aacctcatga ctccggcctg gcttgccacc aaataactta 44535
aaaataatga acttgaaatt atacttcctg aatggagggt tccagatctc cccattttatc 44595
ttgtatggcg tcatcgtcag tattattctc ctttatttca acgctttctg tcttttattg 44655
aagataaatg gaataatcgc ccacaaattg attttctgaa tgatgattaa cccgtttgga 44715
atggttttga tacgttctctg acttaaacc acatgatgac tgaattgagg catcgagata 44775
tgcgactggg cagccagtcg tcttttgacg atgcccata caaaacatga tgccgactga 44835
cggaaatgat aatacgcgga aacaggacgg ggctgttttt gggcagccgg aagttaagcc 44895
cataccagaa acgttgacgt gtactgaaaa atggcgccag gttgcacctg ttcaaagatt 44955
ttctgaaggc gcaggagtat tcattactga tctctccat tgcgccttcg ggaaccaca 45015
ggaccagcta ttttaccgat agtggtttaa aggcgtaagt aatgccgagc atgaagtcac 45075
tggaggcagc ctttgtgtct gcatacataag cggtatgttc atcaccagca tagtgatttt 45135
ttgaaatgct tactttgcca gcattaatgt atttataact ggcgtcaatc ataattattat 45195
ctgttacagc atattttgca ccgataacct cgccccaggc aaagttattt tttgaagcag 45255
acagagtttc attaatacca aaaccaacag gaatggtgtt attacttagc ttcacatgag 45315

cgaggccaac	gcctgcgctg	atatagggag	taaatgccgt	actattgtga	aaatcataat	45375
agccattaac	catgtaagt	gtcattcgga	cctgatTTTT	tacatttatg	tgtactggat	45435
caccaaattgc	aataatatcc	tgcccgctt	tagcatccgt	ctcacctctg	aaagtgggat	45495
ccagttctaa	acgtactgga	agctggaatg	gatacataaaa	gtcataaaccg	atagcaaccc	45555
cgcgcgcaaa	aacgcctttg	gtacggtcag	gtaacgttgc	atgaccatta	actatctcat	4561
cctgggtgaa	ggttgagttg	attccataga	cattgactac	ggatgtcccc	gctttcccg	45675
tgatatatagat	cccttctttt	gctgatgcag	tagcggacca	ggctaccaca	aggggaatga	45735
tgcagactgc	gaaaaagttt	ttcatttcag	aacctgcctt	aatattgggc	taaaagacaa	45795
gtttcacggg	ataggggtgtg	atataacgat	tacataaacg	aagcccaaaa	aacgggtctat	45855
tgtaacgctg	ggttttctgt	aagcgggtaa	aaaatgagat	gaagatttta	aataacaata	45915
cgataatcgt	cggtatggaa	atccatctcc	tcgccaaatt	gccccacgta	cggtttcact	45975
tctacgttat	gtaacgggta	gtgtgagatg	gagcgatgct	gtaagaaaaa	gatgaagatg	46035
aatttgtacc	cgacctggat	aaagcccgtt	atcccgggat	aacgggcaaa	aatatttact	46095
caagtgcctg	ggcgagatct	tgttgtacct	gttgacgctg	ttctgggtgtt	aagactttgc	46155
ttaaatacaa	ataatattta	acccgataat	agcgagcctg	ttgttctatg	ttactgaagg	46215
ctgcaagctg	ctgtttttacg	gcggcgctcat	cccattttacc	ggattttaatc	acctctatca	46275
gcgcaccgtc	tttaattccc	ttcatagaaa	tctgactgac	gtcgggtttcc	agttgttggt	46335
gaagtttttt	gatccgggta	atctgatcgt	ttgtcagctt	cagatgctgg	acaataggat	46395
cctggggcggg	caggggagga	ttggggacag	cggtggcgaa	agcgccaaaa	gaaacgccc	46455
ccagagtcgc	tgccagtaaa	gttgtgcgta	caaagttttt	catgaagata	tcctgataag	46515
ggagtgatta	accgttttta	ttaccacga	atggcgagca	attatcttag	agcctatccc	46575
agtagggcta	ttttacttgc	cattttggac	ctgggcagtg	ctcgccaaaa	cgcgttagcg	46635
ttttgaacgc	cgctagcggc	ggcccgaagg	gcgagcgtag	cgagtcaaac	ctcacgtact	46695
acgtgtacgc	tccggttttt	gcgcgctgtc	cgtgtccaaa	ctggctgcgc	caataacgcc	46755
tggtgggata	ggctcttagt	cagaatacgt	tgcttgcac	attacgccac	gcgaatttgt	46815
tttacggaga	gttacggagt	gaaacaatcc	cgcgcgggtg	agcggcaggt	tgctt	46870

<210> 2

<211> 166

<212> PRT

<213> Salmonella typhimurium

<400> 2

Met Lys Ser Ile Lys Lys Leu Ile Ile Ala Ser Ala Leu Ser Met Met

WO 00/73336

PCT/SE00/01079

31

1 5 10 15
 Ala Ala Ser Cys Tyr Ala Gly Ser Phe Leu Pro Asn Ser Glu Gln Gln
 20 25 30
 Lys Ser Val Asp Ile Val Phe Ser Ser Pro Gln Asp Leu Thr Val Ser
 35 40 45
 Leu Ile Pro Val Ser Gly Leu Lys Ala Gly Lys Asn Ala Pro Ser Ala
 50 55 60
 Lys Ile Ala Lys Leu Val Val Asn Ser Thr Thr Leu Lys Glu Phe Gly
 65 70 75 80
 Val Arg Gly Ile Ser Asn Asn Val Val Asp Ser Thr Gly Thr Ala Trp
 85 90 95
 Arg Val Ala Gly Lys Asn Thr Gly Lys Glu Ile Gly Val Gly Leu Ser
 100 105 110
 Ser Asp Ser Leu Arg Arg Ser Asp Ser Thr Glu Lys Trp Asn Gly Val
 115 120 125
 Asn Trp Met Thr Phe Asn Ser Asn Asp Thr Leu Asp Ile Val Leu Thr
 130 135 140
 Gly Pro Ala Gln Asn Val Thr Ala Asp Thr Tyr Pro Ile Thr Leu Asp
 145 150 155 160
 Val Val Gly Tyr Gln Pro
 165

<210> 3
 <211> 245
 <212> PRT
 <213> Salmonella typhimurium

<400> 3
 Met Lys Ile Val Asn Phe Ala Val Met Ala Val Ala Leu Phe Ala Thr
 1 5 10 15
 Asn Ser Met Val Ser Val Tyr Ala Val Asn Gln Gln Leu Asn Ser Ala
 20 25 30
 Thr Lys Leu Phe Ser Val Lys Leu Gly Ala Thr Arg Val Ile Tyr His
 35 40 45
 Ala Gly Thr Ala Gly Ala Thr Leu Ser Val Ser Asn Pro Gln Asn Tyr
 50 55 60
 Pro Ile Leu Val Gln Ser Ser Val Lys Ala Ala Asp Lys Ser Ser Pro
 65 70 75 80
 Ala Pro Phe Leu Val Met Pro Pro Leu Phe Arg Leu Glu Ala Asn Gln
 85 90 95
 Gln Ser Gln Leu Arg Ile Val Arg Thr Gly Gly Asp Met Pro Thr Asp
 100 105 110
 Arg Glu Thr Leu Gln Trp Val Cys Ile Lys Ala Val Pro Pro Glu Asn

WO 00/73336

PCT/SE00/01079

32

115	120	125
Glu Pro Ser Asp Thr Gln Ala Lys Gly Ala Thr Leu Asp Leu Asn Leu		
130	135	140
Ser Ile Asn Ala Cys Asp Lys Leu Ile Phe Arg Pro Asp Ala Val Lys		
145	150	155 160
Gly Thr Pro Glu Asp Val Ala Gly Asn Leu Arg Trp Val Glu Thr Gly		
	165	170 175
Asn Lys Leu Lys Val Glu Asn Pro Thr Pro Phe Tyr Met Asn Leu Ala		
	180	185 190
Ser Val Thr Val Gly Gly Lys Pro Ile Thr Gly Leu Glu Tyr Val Pro		
	195	200 205
Pro Phe Ala Asp Lys Thr Leu Asn Met Pro Gly Ser Ala His Gly Asp		
	210	215 220
Ile Glu Trp Arg Val Ile Thr Asp Phe Gly Gly Glu Ser His Pro Phe		
225	230	235 240
His Tyr Val Leu Lys		
	245	

<210> 4
 <211> 836
 <212> PRT
 <213> Salmonella typhimurium

<400> 4
 Met Lys Phe Lys Gln Pro Ala Leu Leu Leu Phe Ile Ala Gly Val Val
 1 5 10 15

His Cys Ala Asn Ala His Thr Tyr Thr Phe Asp Ala Ser Met Leu Gly
 20 25 30

Asp Ala Ala Lys Gly Val Asp Met Ser Leu Phe Asn Gln Gly Leu Gln
 35 40 45

Gln Pro Gly Thr Tyr Arg Val Asp Val Met Val Asn Gly Lys Arg Val
 50 55 60

Asp Thr Arg Asp Val Val Phe Lys Leu Glu Lys Asp Gly Gln Gly Thr
 65 70 75 80

Pro Val Leu Ala Pro Cys Leu Thr Val Ser Gln Leu Ser Arg Tyr Gly
 85 90 95

Val Lys Thr Glu Asp Tyr Pro Gln Leu Trp Lys Ala Ala Lys Pro Pro
 100 105 110

Asp Glu Cys Ala Asp Leu Thr Ala Ile Pro Gln Ala Lys Ala Val Leu
 115 120 125

Asp Ile Asn Asn Gln Gln Leu Gln Leu Ser Ile Pro Gln Leu Ala Leu
 130 135 140

Arg Pro Glu Phe Lys Gly Ile Ala Pro Glu Asp Leu Trp Asp Asp Gly

WO 00/73336

PCT/SE00/01079

33

145	150	155	160
Ile Pro Ala Phe	Leu Met Asn Tyr Ser	Ala Arg Thr Thr	Gln Thr Asp
	165	170	175
Tyr Lys Met Asp	Met Val Gly Arg Asp	Asn Ser Ser Trp	Val Gln Leu
	180	185	190
Gln Pro Gly Ile	Asn Ile Gly Ala Trp	Arg Val Arg Asn	Ala Thr Ser
	195	200	205
Trp Gln Arg Ser	Ser Gln Leu Ser Gly	Lys Trp Gln Ala	Ala Tyr Thr
	210	215	220
Tyr Ala Glu Arg	Gly Leu Tyr Ser Leu	Lys Ser Arg Leu	Thr Leu Gly
	225	230	235
Gln Lys Thr Ser	Gln Gly Glu Ile Phe	Asp Ser Val Pro	Phe Thr Gly
	245	250	255
Val Met Leu Ala	Ser Asp Asp Asn Met	Val Pro Tyr Ser	Glu Arg Gln
	260	265	270
Phe Ala Pro Val	Val Arg Gly Ile Ala	Arg Thr Gln Ala	Arg Val Glu
	275	280	285
Val Lys Gln Asn	Gly Tyr Thr Ile Tyr	Asn Thr Thr Val	Ala Pro Gly
	290	295	300
Pro Phe Ala Leu	Arg Asp Leu Ser Val	Thr Asp Ser Ser	Gly Asp Leu
	305	310	315
His Val Thr Val	Trp Glu Ala Asp Gly	Ser Thr Gln Met	Phe Val Val
	325	330	335
Pro Tyr Gln Thr	Pro Ala Ile Ala Leu	His Gln Gly Tyr	Leu Lys Tyr
	340	345	350
Ser Leu Leu Ala	Gly Arg Tyr Arg Ser	Ser Asp Ser Ala	Thr Asp Lys
	355	360	365
Arg Gln Ile Ala	Gln Ala Thr Leu Met	Tyr Gly Leu Pro	Trp Asn Leu
	370	375	380
Thr Ala Tyr Gly	Gly Ile Gln Ser Ala	Thr His Asn Gln	Ala Ala Leu
	385	390	395
Leu Gly Leu Gly	Gly Ser Leu Gly Arg	Trp Gly Ser Leu	Ser Val Asp
	405	410	415
Gly Ser Asp Thr	His Ser Gln Arg Gln	Gly Glu Ala Val	Gln Gln Gly
	420	425	430
Ala Ser Trp Arg	Leu Arg Tyr Ser Asn	Gln Leu Thr Ala	Thr Gly Thr
	435	440	445
Asn Phe Phe Leu	Thr Arg Trp Gln Tyr	Ala Ser Gln Gly	Tyr Asn Thr
	450	455	460
Leu Ser Asp Val	Leu Asp Ser Tyr Arg	His Asn Gly Asn	Arg Leu Trp
	465	470	475
			480

WO 00/73336

PCT/SE00/01079

34

Ser Trp Arg Glu Asn Leu Gln Pro Ser Ser Arg Thr Thr Leu Met Leu
 485 490 495

Ser Gln Ser Trp Gly Arg His Leu Gly Asn Leu Ser Leu Thr Gly Ser
 500 505 510

Arg Thr Asp Trp Arg Asn Arg Pro Gly His Asp Asp Ser Tyr Gly Leu
 515 520 525

Ser Trp Gly Thr Ser Ile Gly Gly Gly Ser Leu Ser Leu Asn Trp Asn
 530 535 540

Gln Asn Arg Thr Leu Trp Arg Asn Gly Ala His Arg Lys Glu Asn Ile
 545 550 555 560

Thr Ser Leu Trp Phe Ser Met Pro Leu Ser Arg Trp Thr Gly Asn Asn
 565 570 575

Val Ser Ala Ser Trp Gln Met Thr Ser Pro Ser His Gly Gly Gln Thr
 580 585 590

Gln Gln Val Gly Val Asn Gly Glu Ala Phe Ser Gln Gln Leu Asp Trp
 595 600 605

Glu Val Arg Gln Ser Tyr Arg Ala Asp Ala Pro Pro Gly Gly Gly Asn
 610 615 620

Asn Ser Ala Leu His Leu Ala Trp Asn Gly Asp Tyr Gly Leu Leu Gly
 625 630 635 640

Gly Asp Tyr Ser Tyr Ser Arg Ala Met Arg Gln Met Gly Val Asn Ile
 645 650 655

Ala Gly Gly Ile Val Ile His His His Gly Val Thr Leu Gly Gln Pro
 660 665 670

Leu Gln Gly Ser Val Ala Leu Val Glu Ala Pro Gly Ala Ser Gly Val
 675 680 685

Pro Val Gly Gly Trp Pro Gly Val Lys Thr Asp Phe Arg Gly Asp Thr
 690 695 700

Thr Val Gly Asn Leu Asn Val Tyr Gln Glu Asn Thr Val Ser Leu Asp
 705 710 715 720

Pro Ser Arg Leu Pro Asp Asp Ala Glu Val Thr Gln Thr Asp Val Arg
 725 730 735

Val Val Pro Thr Glu Gly Ala Val Val Glu Ala Lys Phe His Thr Arg
 740 745 750

Ile Gly Ala Arg Ala Leu Met Thr Leu Lys Arg Glu Asp Gly Ser Ala
 755 760 765

Ile Pro Phe Gly Ala Gln Val Thr Val Asn Gly Gln Asp Gly Ser Ala
 770 775 780

Ala Leu Val Asp Thr Asp Ser Gln Val Tyr Leu Thr Gly Leu Ala Asp
 785 790 795 800

35

Lys Gly Glu Leu Thr Val Lys Trp Gly Ala Gln Gln Cys Arg Val Asn
805 810 815

Tyr Arg Leu Pro Ala His Lys Gly Ile Ala Gly Leu Tyr Gln Met Ser
820 825 830

Gly Leu Cys Arg
835

<210> 5

<211> 156

<212> PRT

<213> Salmonella typhimurium

<400> 5

Met Trp Met Lys Ile Gln Arg Val Lys Thr Val Ile Tyr Ser Val Ser
1 5 10 15

Leu Leu Val Ala Ala Ser Ser Leu Val Pro Ile Ala Asn Ala Ala Glu
20 25 30

Lys Leu Gln Thr Thr Leu Arg Val Gly Thr Tyr Phe Arg Ala Gly His
35 40 45

Val Pro Asp Gly Met Val Leu Ala Gln Gly Trp Val Thr Tyr His Gly
50 55 60

Ser His Ser Gly Phe Arg Val Trp Ser Asp Glu Gln Lys Ala Gly Asn
65 70 75 80

Thr Pro Thr Val Leu Leu Leu Ser Gly Gln Gln Asp Pro Arg His His
85 90 95

Ile Gln Val Arg Leu Glu Gly Glu Gly Trp Gln Pro Asp Thr Val Ser
100 105 110

Gly Arg Gly Ala Ile Leu Arg Thr Ala Ala Asp Asn Ala Ser Phe Ser
115 120 125

Val Val Val Asp Gly Asn Gln Glu Val Pro Ala Asp Thr Trp Thr Leu
130 135 140

Asp Phe Lys Ala Cys Ala Leu Ala Gln Glu Asp Thr
145 150 155

SEQUENCE LISTING NO. 2

<110> Folkesson, Anders

<120> The complete sequence of the tcf insert of Salmonella enterica serovar Typhi.

<130> The tcf insert in Salmonella typhi

<140>

<141>

<160> 6

<170> PatentIn Ver. 2.1

<210> 1

<211> 9253

<212> DNA

<213> Salmonella typhi

<220>

<221> CDS

<222> (1898)..(2608)

<223> tcfA putative fimbrial subunit

<220>

<221> CDS

<222> (2659)..(3234)

<223> tcfB putative fimbrial subunit

<220>

<221> CDS

<222> (3360)..(6029)

<223> tcfC putative fimbrial subunit

<220>

<221> CDS

<222> (6052)..(7131)

<223> tcfD putative fimbrial subunit

<220>

<221> CDS

<222> (7264)..(7719)

<223> tinR putative transcriptional regulator

<400> 1

tgtaagtatc ccgcataatc gagccattca catttagaga tcatccggca taatcaatct 60

gccaacgcag gagatcgctg tgcgtaaagc ccgtattact gcgcaccaga tcatcgctgt 120

gattagatca gttgaatccg gacggactgt taaagatgtc taccgggagg ccggtatttc 180

tgaagccacc agggacaact ggaagtctgg atacggcgcc atggaagctt ctgatattaa 240

atcttgagga tgtcaacgcc aggatttatg gtcgtttttc gctattttta atatccgctg 300

tttgatcact tctgctgtcc gctttccgcc atttcattct cactgattgg cgttgcgctt 360

ttggtcagcg ccccgacttt gcgtttttcc ttcagttggt aactttctcc tttgatattc 420

cgc gaa gag gcc acc tcg ctg ttc gtc tat tct aaa tca gat cat gtg 2059
 Arg Glu Glu Ala Thr Ser Leu Phe Val Tyr Ser Lys Ser Asp His Val
 40 45 50

caa tat att cga aca aga atc atg cgt att gaa cac ccc ggt atg cca 2107
 Gln Tyr Ile Arg Thr Arg Ile Met Arg Ile Glu His Pro Gly Met Pro
 55 60 65 70

cag gag aag gag gta cca gca ggg aat gat ata gag aca gga ctt gtt 2155
 Gln Glu Lys Glu Val Pro Ala Gly Asn Asp Ile Glu Thr Gly Leu Val
 75 80 85

gtc tcc ccg gag aaa ttt gct ctt tcc ccg gga aca aaa aaa aca ata 2203
 Val Ser Pro Glu Lys Phe Ala Leu Ser Pro Gly Thr Lys Lys Thr Ile
 90 95 100

cgt gtt atc agt act cag gca ccg gaa aga gag gaa gcc tgg cgg gta 2251
 Arg Val Ile Ser Thr Gln Ala Pro Glu Arg Glu Glu Ala Trp Arg Val
 105 110 115

tac ttc gag gct gtt cct gaa ctg gaa gat gat cca cag gca ggc gga 2299
 Tyr Phe Glu Ala Val Pro Glu Leu Glu Asp Asp Pro Gln Ala Gly Gly
 120 125 130

aag caa aat tca tcc gta agt gtg aat ctt gtc tgg ggg gtg ttg ctg 2347
 Lys Gln Asn Ser Ser Val Ser Val Asn Leu Val Trp Gly Val Leu Leu
 135 140 145 150

cgt gtt tct ccg tca gac ccc agg cct gcg ctg gta acg gac ggt cac 2395
 Arg Val Ser Pro Ser Asp Pro Arg Pro Ala Leu Val Thr Asp Gly His
 155 160 165

cac ctg ctg aat acg gga aac aca cgg ctt tct ctt att cgg gct ggc 2443
 His Leu Leu Asn Thr Gly Asn Thr Arg Leu Ser Leu Ile Arg Ala Gly
 170 175 180

aac tgc gac acc aca tgc cac tgg cag aat ata ggc aaa agt att tat 2491
 Asn Cys Asp Thr Thr Cys His Trp Gln Asn Ile Gly Lys Ser Ile Tyr
 185 190 195

ccc ggc ggg agt gct gat att ccg gcc gga ata aaa agt aat gca ttt 2539
 Pro Gly Gly Ser Ala Asp Ile Pro Ala Gly Ile Lys Ser Asn Ala Phe
 200 205 210

cgt gtg gaa tat cgt acg ggt gca aat tca ccg gta atc tct gct gat 2587
 Arg Val Glu Tyr Arg Thr Gly Ala Asn Ser Pro Val Ile Ser Ala Asp
 215 220 225 230

tta aca gca gcc gga aag taa aaacacacgg agcgtacgct ataccctaca 2638
 Leu Thr Ala Ala Gly Lys
 235

tttattctca gggggagcgg atg tat acc gag tgt aca tat atc act gta ata 2691
 Met Tyr Thr Glu Cys Thr Tyr Ile Thr Val Ile
 240 245

aac aac aaa gca agg tta ttt ttt atg aac atg aaa aca tct ttt att 2739
 Asn Asn Lys Ala Arg Leu Phe Phe Met Asn Met Lys Thr Ser Phe Ile

PCT/SE00/01079

41

acc	cga	aat	gcc	cgt	att	gat	att	tat	cgt	ggc	agc	gag	ttg	ctg	ggg	4250
Thr	Arg	Asn	Ala	Arg	Ile	Asp	Ile	Tyr	Arg	Gly	Ser	Glu	Leu	Leu	Gly	
				715					720					725		
agt	cag	ttc	ctg	acc	ccg	gga	atg	cat	acc	ctg	gat	act	cat	tct	ctt	4298
Ser	Gln	Phe	Leu	Thr	Pro	Gly	Met	His	Thr	Leu	Asp	Thr	His	Ser	Leu	
				730					735					740		
cca	ccg	gga	agc	tat	cct	ctg	gcg	ttg	cgg	gtg	tat	gag	gat	ggg	att	4346
Pro	Pro	Gly	Ser	Tyr	Pro	Leu	Ala	Leu	Arg	Val	Tyr	Glu	Asp	Gly	Ile	
				745					750					755		
ctg	cgg	cga	acg	gag	acc	cag	ccc	ttc	agt	aag	ggg	ggc	aat	agc	ttc	4394
Leu	Arg	Arg	Thr	Glu	Thr	Gln	Pro	Phe	Ser	Lys	Gly	Gly	Asn	Ser	Phe	
				760					765					770		
agt	gca	cag	acc	cag	tgg	ttt	att	cag	ggc	ggg	ctg	gaa	gat	acc	ggg	4442
Ser	Ala	Gln	Thr	Gln	Trp	Phe	Ile	Gln	Gly	Gly	Leu	Glu	Asp	Thr	Gly	
				775					780					785		
gat	aaa	gcc	agc	cat	tat	gac	ggc	gag	act	gtc	atg	gct	gcc	gga	ttc	4490
Asp	Lys	Ala	Ser	His	Tyr	Asp	Gly	Glu	Thr	Val	Met	Ala	Ala	Gly	Phe	
				795					800					805		
caa	act	ggg	ctg	cgg	aaa	aat	atc	agt	ctg	acc	gaa	ggc	atc	tct	ctg	4538
Gln	Thr	Gly	Leu	Arg	Lys	Asn	Ile	Ser	Leu	Thr	Glu	Gly	Ile	Ser	Leu	
				810					815					820		
gca	cat	gag	gcc	tgg	tac	agt	gaa	acc	cga	ctg	aat	tca	cag	cat	gca	4586
Ala	His	Glu	Ala	Trp	Tyr	Ser	Glu	Thr	Arg	Leu	Asn	Ser	Gln	His	Ala	
				825					830					835		
gtg	ctg	gat	ggc	acg	ctg	gac	ctt	tct	gcc	ggg	ata	ctg	cat	ggg	aca	4634
Val	Leu	Asp	Gly	Thr	Leu	Asp	Leu	Ser	Ala	Gly	Ile	Leu	His	Gly	Thr	
				840					845					850		
gac	agc	acg	agc	ggc	aac	act	gag	cag	gtg	aca	tac	aac	gac	gga	ttt	4682
Asp	Ser	Thr	Ser	Gly	Asn	Thr	Glu	Gln	Val	Thr	Tyr	Asn	Asp	Gly	Phe	
				855					860					865		
tcc	gcg	agt	ctg	tgg	cgt	aac	cat	acg	gaa	agt	gat	gcc	tgt	agt	ggc	4730
Ser	Ala	Ser	Leu	Trp	Arg	Asn	His	Thr	Glu	Ser	Asp	Ala	Cys	Ser	Gly	
				875					880					885		
cgt	cat	cca	cag	tca	gtg	cat	gcc	agt	atg	acc	tgc	cag	act	tcg	atg	4778
Arg	His	Pro	Gln	Ser	Val	His	Ala	Ser	Met	Thr	Cys	Gln	Thr	Ser	Met	
				890					895					900		
aac	gcc	tcc	ctg	tcg	gtt	tcg	gtg	ggg	aac	tgg	tat	gcc	cta	ctg	gga	4826
Asn	Ala	Ser	Leu	Ser	Val	Ser	Val	Gly	Asn	Trp	Tyr	Ala	Leu	Leu	Gly	
				905					910					915		
tac	agt	acc	agc	agg	aca	gaa	ggc	cgg	ccg	gtt	tac	cgg	gga	tat	gat	4874
Tyr	Ser	Thr	Ser	Arg	Thr	Glu	Gly	Arg	Pro	Val	Tyr	Arg	Gly	Tyr	Asp	
				920					925					930		
gat	aac	agt	gac	aaa	gaa	aat	gtg	ttc	tgg	cga	cag	gca	tac	atc	cct	4922
Asp	Asn	Ser	Asp	Lys	Glu	Asn										

PCT/SE00/01079

42

gcc	tct	cac	cgc	gaa	tct	gct	cag	gct	agt	gca	acg	tac	agc	ctt	aat	4970
Ala	Ser	His	Arg	Glu	Ser	Ala	Gln	Ala	Ser	Ala	Thr	Tyr	Ser	Leu	Asn	
				955					960				965			
atg	gct	ggc	atg	aat	att	aat	acc	cat	ggg	gga	gta	tgg	cga	acc	cga	5018
Met	Ala	Gly	Met	Asn	Ile	Asn	Thr	His	Gly	Gly	Val	Trp	Arg	Thr	Arg	
				970					975				980			
aat	gac	gga	gtg	aat	gat	gat	ggc	ttg	ttt	atg	agt	gtc	agt	gtg	tca	5066
Asn	Asp	Gly	Val	Asn	Asp	Asp	Gly	Leu	Phe	Met	Ser	Val	Ser	Val	Ser	
				985					990				995			
tat	gcc	tct	caa	cca	ccg	aca	atg	act	ggc	agt	aat	agg	tat	acc	tca	5114
Tyr	Ala	Ser	Gln	Pro	Pro	Thr	Met	Thr	Gly	Ser	Asn	Arg	Tyr	Thr	Ser	
1000				1005				1010								
gcc	ggg	acc	gat	att	cac	agt	agc	cgg	aat	caa	aaa	aca	cag	acg	tcc	5162
Ala	Gly	Thr	Asp	Ile	His	Ser	Ser	Arg	Asn	Gln	Lys	Thr	Gln	Thr	Ser	
1015				1020				1025				1030				
tgg	aat	gtg	aac	cat	gtg	aga	tcc	tgg	cag	cag	gat	ctg	tat	cgt	gaa	5210
Trp	Asn	Val	Asn	His	Val	Arg	Ser	Trp	Gln	Gln	Asp	Leu	Tyr	Arg	Glu	
				1035					1040				1045			
ctg	tcg	gtg	ggc	ttc	tcc	ggc	tat	aac	gac	gac	agc	tgg	agc	ggg	agt	5258
Leu	Ser	Val	Gly	Phe	Ser	Gly	Tyr	Asn	Asp	Asp	Ser	Trp	Ser	Gly	Ser	
1050								1055				1060				
ctc	ggc	gga	cgc	atg	agc	ggc	cgt	atg	ggc	gaa	ctg	agc	gcc	act	atc	5306
Leu	Gly	Gly	Arg	Met	Ser	Gly	Arg	Met	Gly	Glu	Leu	Ser	Ala	Thr	Ile	
1065								1070				1075				
agt	aac	tcc	cat	caa	cgt	aat	gcg	ggc	agc	gcc	agt	tca	ctc	acc	gct	5354
Ser	Asn	Ser	His	Gln	Arg	Asn	Ala	Gly	Ser	Ala	Ser	Ser	Leu	Thr	Ala	
1080				1085				1090								
ggc	tac	agc	tcg	tct	ctg	gcg	tta	tcc	cgt	aat	gga	ctg	ttc	tgg	gga	5402
Gly	Tyr	Ser	Ser	Ser	Leu	Ala	Leu	Ser	Arg	Asn	Gly	Leu	Phe	Trp	Gly	
1095				1100				1105				1110				
ggc	ggc	cag	gac	ggc	gaa	ccg	gcc	tct	ggc	atg	gcg	gtg	aac	gtg	gag	5450
Gly	Gly	Gln	Asp	Gly	Glu	Pro	Ala	Ser	Gly	Met	Ala	Val	Asn	Val	Glu	
				1115					1120				1125			
tca	gag	ggg	gac	gag	ggc	agt	agc	ggg	aaa	gta	gtc	agc	gtt	cgt	ggc	5498
Ser	Glu	Gly	Asp	Glu	Gly	Ser	Ser	Gly	Lys	Val	Val	Ser	Val	Arg	Gly	
				1130					1135				1140			
agc	agc	cag	ccg	ttc	agt	ctc	ggc	ttt	ggc	cag	cag	tcg	ctg	ttg	ctg	5546
Ser	Ser	Gln	Pro	Phe	Ser	Leu	Gly	Phe	Gly	Gln	Gln	Ser	Leu	Leu	Leu	
1145				1150				1155								
atg	gaa	ggc	tat	aac	gcc	acg	gag	gtg	acc	att	gag	gat	gca	ggg	gtt	5594
Met	Glu	Gly	Tyr	Asn	Ala	Thr	Glu	Val	Thr	Ile	Glu	Asp	Ala	Gly	Val	
1160				1165				1170								
agt	tca	cag	ggc	atg	gca	ggc	gta	aaa	gcg	gga	ggg	gga	agc	agg	tgt	5642
Ser	Ser	Gln	Gly	Met	Ala	Gly	Val	Lys	Ala	Gly	Gly	Gly	Ser	Arg	Cys	

WO 00/73336

PCT/SE00/01079

43

1175	1180	1185	1190	
tac ttc ctg aca ccc ggg cat ctg ctg gtt cac aac atc agc gcc agt				5690
Tyr Phe Leu Thr Pro Gly His Leu Leu Val His Asn Ile Ser Ala Ser				
	1195	1200	1205	
atg agc cga ctg tac gtt ggc cgc gta ctg gac aag gat ggc aga ccg				5738
Met Ser Arg Leu Tyr Val Gly Arg Val Leu Asp Lys Asp Gly Arg Pro				
	1210	1215	1220	
ctg ctg gac gca cag cca ctg aac tat cca ttt ttg tgc ttg gga cct				5786
Leu Leu Asp Ala Gln Pro Leu Asn Tyr Pro Phe Leu Ser Leu Gly Pro				
	1225	1230	1235	
tcc ggg cga ttt agc ctg cag agc gag cat aaa gaa tcc agc ctg tgg				5834
Ser Gly Arg Phe Ser Leu Gln Ser Glu His Lys Glu Ser Ser Leu Trp				
	1240	1245	1250	
ctg ctg tct aaa aac agg atc ctg cgt tgt ccg atg tca gta cat aaa				5882
Leu Leu Ser Lys Asn Arg Ile Leu Arg Cys Pro Met Ser Val His Lys				
	1255	1260	1265	1270
cgt cgg gat gtt atg cag gta gtc ggt gat gtg cgg tgt gaa tta agt				5930
Arg Arg Asp Val Met Gln Val Val Gly Asp Val Arg Cys Glu Leu Ser				
	1275	1280	1285	
gac gtg gat gcc ctg cca cag gcg ttg caa ata tgc ccg cgg gtc atc				5978
Asp Val Asp Ala Leu Pro Gln Ala Leu Gln Ile Ser Pro Arg Val Ile				
	1290	1295	1300	
cgt ttg ctg aac gtg gca ggt ttg ctg cgc cat tcc gtt cag gaa gcc				6026
Arg Leu Leu Asn Val Ala Gly Leu Leu Arg His Ser Val Gln Glu Ala				
	1305	1310	1315	
tga cgtagagata aaggcgtaa ct atg agt aat aaa atg aag tgg acg agt				6078
Met Ser Asn Lys Met Lys Trp Thr Ser				
	1320	1325		
atg aca gcc cat tgg tca gca att att aat ttc atc cga aaa tat gtt				6126
Met Thr Ala His Trp Ser Ala Ile Ile Asn Phe Ile Arg Lys Tyr Val				
	1330	1335	1340	
tat cca gca agg ata att gcc atc ctg ctg atg gct ggc gct aca ctg				6174
Tyr Pro Ala Arg Ile Ile Ala Ile Leu Leu Met Ala Gly Ala Thr Leu				
	1345	1350	1355	1360
cca caa gtc gcc gat gcg att acc gtc gac ctg aat tac gac aag aac				6222
Pro Gln Val Ala Asp Ala Ile Thr Val Asp Leu Asn Tyr Asp Lys Asn				
	1365	1370	1375	
aat gta gcg gtc atc act cct gtc tgg tcc caa gaa tgg agt gta gca				6270
Asn Val Ala Val Ile Thr Pro Val Trp Ser Gln Glu Trp Ser Val Ala				
	1380	1385	1390	
aat gtg ttg ggg gga tgg gta tgt cgt tca aac agg aat gaa aat gag				6318
Asn Val Leu Gly Gly Trp Val Cys Arg Ser Asn Arg Asn Glu Asn Glu				
	1395	1400	1405	
ggg gcg tgt gaa gaa aca cat ttg gta tgg tgg tat gct ttt gga gct				6366
Gly Ala Cys Glu Glu Thr His Leu Val Trp Trp Tyr Ala Phe Gly Ala				

1410	1415	1420	
tat tca aaa att cgt ctg cgt ttc aga gaa caa atc agc cat gcc gaa			6414
Tyr Ser Lys Ile Arg Leu Arg Phe Arg Glu Gln Ile Ser His Ala Glu			
1425	1430	1435	1440
att acg ctc ata ctg ctc ggc agt gtt cgt gat gcc tgt tat act ggt			6462
Ile Thr Leu Ile Leu Leu Gly Ser Val Arg Asp Ala Cys Tyr Thr Gly			
1445	1450	1455	
gtc atc aac atg aac gct gct gca tgt caa tgg ggt agg tcg ctg aaa			6510
Val Ile Asn Met Asn Ala Ala Ala Cys Gln Trp Gly Arg Ser Leu Lys			
1460	1465	1470	
ctt agg ata cct tca gaa gag ctt gcg aag ata cct aca agc gga aca			6558
Leu Arg Ile Pro Ser Glu Glu Leu Ala Lys Ile Pro Thr Ser Gly Thr			
1475	1480	1485	
tgg aaa gca acg tta gtc ctg gat tat tta caa tgg ggc gga gac gat			6606
Trp Lys Ala Thr Leu Val Leu Asp Tyr Leu Gln Trp Gly Gly Asp Asp			
1490	1495	1500	
cct tta ggc aca tca act aca gat atc acg ctg aat gta aca gac cac			6654
Pro Leu Gly Thr Ser Thr Thr Asp Ile Thr Leu Asn Val Thr Asp His			
1505	1510	1515	1520
ttt gct gaa aat gcg gct att tac ttt ccg caa ttt ggt aca gca acg			6702
Phe Ala Glu Asn Ala Ala Ile Tyr Phe Pro Gln Phe Gly Thr Ala Thr			
1525	1530	1535	
ccc cgg gtg gac ctg aat ctt cac cgg atg aat gcc tca caa atg tcg			6750
Pro Arg Val Asp Leu Asn Leu His Asp Met Asn Ala Ser Gln Met Ser			
1540	1545	1550	
ggc agg gct aat ctg gat atg tgt ctg tat gac gga ggt gtg aaa gcc			6798
Gly Arg Ala Asn Leu Asp Met Cys Leu Tyr Asp Gly Gly Val Lys Ala			
1555	1560	1565	
cgt tca tta cag atg aag ata gaa gga agc aat aag tca ggt acg gga			6846
Arg Ser Leu Gln Met Lys Ile Glu Gly Ser Asn Lys Ser Gly Thr Gly			
1570	1575	1580	
ttt cag gtt ata aag agc gat tct gct gat acg att gat tat gcg gtc			6894
Phe Gln Val Ile Lys Ser Asp Ser Ala Asp Thr Ile Asp Tyr Ala Val			
1585	1590	1595	1600
agt atg aat tat ggg gga cga agt att cct gtc acc cgt ggc gtg gag			6942
Ser Met Asn Tyr Gly Gly Arg Ser Ile Pro Val Thr Arg Gly Val Glu			
1605	1610	1615	
ttc agt ctg gat aac gtg gat aaa gca gca acg cgt ccg gtg gta ctt			6990
Phe Ser Leu Asp Asn Val Asp Lys Ala Ala Thr Arg Pro Val Val Leu			
1620	1625	1630	
ccc ggg caa cgg cag gcg gta cgt tgt gtg cca gtg ccc ctt acc ctg			7038
Pro Gly Gln Arg Gln Ala Val Arg Cys Val Pro Val Pro Leu Thr Leu			
1635	1640	1645	

aca aca caa ccc ttt aac atc aga gag aag cgt tct ggt gag tat cag 7086
 Thr Thr Gln Pro Phe Asn Ile Arg Glu Lys Arg Ser Gly Glu Tyr Gln
 1650 1655 1660

gga acg ctg aca gtg aca atg ctg atg gga aca caa acc ccc tga 7131
 Gly Thr Leu Thr Val Thr Met Leu Met Gly Thr Gln Thr Pro
 1665 1670 1675

cagtaattat ttatttttatt gatattctttc ttatatgggtt ttttaaatca gagttctctt 7191
 tatatacttg tttttatttaa taaagagaat ctattcactt atgaaaatca atgcgtgagg 7251

ttctgctttc ct atg act gtg tat tta gat gat aaa gat aaa gaa tta ttg 7302
 Met Thr Val Tyr Leu Asp Asp Lys Asp Lys Glu Leu Leu
 1680 1685 1690

aaa gaa atc caa aaa gat tgt gca caa act tta tgg caa ctt gca tat 7350
 Lys Glu Ile Gln Lys Asp Cys Ala Gln Thr Leu Trp Gln Leu Ala Tyr
 1695 1700 1705

aaa gtg gga ctt acg ccc aca cca tgt ttc aaa cgt tta aaa aza ctt 7398
 Lys Val Gly Leu Thr Pro Thr Pro Cys Phe Lys Arg Leu Lys Lys Leu
 1710 1715 1720

aaa gac agg ggg gtt atc att ggt cag ttc gct tta ttg gat aag gaa 7446
 Lys Asp Arg Gly Val Ile Ile Gly Gln Phe Ala Leu Leu Asp Lys Glu
 1725 1730 1735 1740

aaa cta ggt ctt tca ctt aat gtc ttt att atg att aac ata tct gag 7494
 Lys Leu Gly Leu Ser Leu Asn Val Phe Ile Met Ile Asn Ile Ser Glu
 1745 1750 1755

gag caa tac gct agt att tct gag aaa ata aag tca atg cct gag gtt 7542
 Glu Gln Tyr Ala Ser Ile Ser Glu Lys Ile Lys Ser Met Pro Glu Val
 1760 1765 1770

att gcc ttt tat cga att tct gga tca ttt aat tat tta atg cat aca 7590
 Ile Ala Phe Tyr Arg Ile Ser Gly Ser Phe Asn Tyr Leu Met His Thr
 1775 1780 1785

gta ttt aca gat atg aac gat tac tat agt ttt tat gag aaa ata ata 7638
 Val Phe Thr Asp Met Asn Asp Tyr Tyr Ser Phe Tyr Glu Lys Ile Ile
 1790 1795 1800

tta act aat tct tca att agt gga tct gca tcg agc ttt gtt ctt gag 7686
 Leu Thr Asn Ser Ser Ile Ser Gly Ser Ala Ser Ser Phe Val Leu Glu
 1805 1810 1815 1820

caa ata aag gaa aca aac gaa ctg tca gtg tga aagtgtgatg tgtacttact 7739
 Gln Ile Lys Glu Thr Asn Glu Leu Ser Val
 1825 1830

gatttaatac attattatcc ttcttacgga acaacaacgg cagattgcgg ctgttgaaca 7799
 aggatttttaa tcagcagtgg tgaaattaag cggcacagaa taacacagcg gaatatcaca 7859
 tggttaaata tcaccccggtg catgtaacaa aaaaccgcat taaaacagat gatgttactg 7919
 atatttattt cggtgaaccc ttctggaaaa aaggcgaaaa ccacataatt gagtcattga 7979

tgtttttga agagttacaa aagtcattta atttattcaa ccataaatat gggttaaata 8039
 aatatatact caggatcccc tgggaatttg tgctcataca tatggaaagg atcagtaaat 8099
 taaatagcgt cgggttattt gctgtttctg ttgactttta taacaaccac aaatttctga 8159
 gcgagtagat caggagtcgc agagattatg gtatggaagt ttgggttgat ttttctggta 8219
 aacattctta ttccagtga attaaaaacc ttggattctt ttttcaggct tgcgtagtgc 8279
 ctctgatcc taattttatt agtagtggtt atcattatca taagttccaa aagattcttg 8339
 tcggggatat aaatgatga gaacagaggg ccgtgtacca gaacgaagtt gattacatgt 8399
 atggaatgca atggccatcg tcatatgacg gttttttctt tcgggatcat aaaaaaatg 8459
 aaacttgggtg tatataacag aaggagtga aatttgaatc aaaaatatct tatttatttt 8519
 ttgtttaatt attgttttgt tttttattac gattaaatat aaagaacatc attgttcgtg 8579
 cgggtggggag gctggaagtt taggggatga ccgtttatca acaattttat tacagccacc 8639
 atacgaatgg ttatatatg cactagatgt attattttag tttaatatat cgatgggtgc 8699
 tatttgcatt gatgatgttc cgttacatta aggaatatac atctgtatct cgttatacgc 8759
 acactcacat tactaatcat tattaatatg agtgtgggtc ttgttttacg catgcatggc 8819
 tgcattgtgac gttaaattta aatgagctga ctgtatgaat tctaaatact ttagagaggc 8879
 gttttttgtc tcggtaggtg ttatattatt attttatttg gtgttatttg cagccagtgc 8939
 tcatgctgaa ggcgggttca gatctggagg cattgggtta tttatgacgg gaacaagaga 8999
 gatgctactg tagagataat aaattctgct aaagattccc caattcttgt gcattgacat 9059
 cctccacgtc ctgaagggcg tgggttcctg ctccaacggg ctgcctgact gcacgctcct 9119
 tccacaggca agcacggcgt gtcccgtctt aaaatgttac gcgcgccgtt tacatcggcg 9179
 ttgcgagtat atcttcatac cagacacttg taagtatctc gcataatcgt gccattcaca 9239
 ttttagagatc atac 9253

<210> 2

<211> 236

<212> PRT

<213> Salmonella typhi

<400> 2

Met Asn Phe Lys Asp Thr Leu Pro Gly Val Phe Leu Cys Val Ala Met
 1 5 10 15

Phe Ala Cys Gly His Ala Arg Ala Asn Met Leu Val Tyr Pro Met Ala
 20 25 30

Ala Glu Ile Asn Ser Ser Arg Glu Glu Ala Thr Ser Leu Phe Val Tyr
 35 40 45

WO 00/73336

PCT/SE00/01079

47

Ser Lys Ser Asp His Val Gln Tyr Ile Arg Thr Arg Ile Met Arg Ile
 50 55 60
 Glu His Pro Gly Met Pro Gln Glu Lys Glu Val Pro Ala Gly Asn Asp
 65 70 75 80
 Ile Glu Thr Gly Leu Val Val Ser Pro Glu Lys Phe Ala Leu Ser Pro
 85 90 95
 Gly Thr Lys Lys Thr Ile Arg Val Ile Ser Thr Gln Ala Pro Glu Arg
 100 105 110
 Glu Glu Ala Trp Arg Val Tyr Phe Glu Ala Val Pro Glu Leu Glu Asp
 115 120 125
 Asp Pro Gln Ala Gly Gly Lys Gln Asn Ser Ser Val Ser Val Asn Leu
 130 135 140
 Val Trp Gly Val Leu Leu Arg Val Ser Pro Ser Asp Pro Arg Pro Ala
 145 150 155 160
 Leu Val Thr Asp Gly His His Leu Leu Asn Thr Gly Asn Thr Arg Leu
 165 170 175
 Ser Leu Ile Arg Ala Gly Asn Cys Asp Thr Thr Cys His Trp Gln Asn
 180 185 190
 Ile Gly Lys Ser Ile Tyr Pro Gly Gly Ser Ala Asp Ile Pro Ala Gly
 195 200 205
 Ile Lys Ser Asn Ala Phe Arg Val Glu Tyr Arg Thr Gly Ala Asn Ser
 210 215 220
 Pro Val Ile Ser Ala Asp Leu Thr Ala Ala Gly Lys
 225 230 235

<210> 3
 <211> 191
 <212> PRT
 <213> Salmonella typhi

<400> 3
 Met Tyr Thr Glu Cys Thr Tyr Ile Thr Val Ile Asn Asn Lys Ala Arg
 1 5 10 15
 Leu Phe Phe Met Asn Met Lys Thr Ser Phe Ile Ala Ala Val Ala
 20 25 30
 Leu Ala Thr Val Tyr Ser Phe Ser Val Ser Ala Val Gln Lys Asp Ile
 35 40 45
 Thr Val Thr Ala Asn Ile Asp Ser Thr Leu Glu Leu Leu Gln Ala Asp
 50 55 60
 Gly Ser Ser Leu Pro Ser Thr Met Lys Leu Asp Phe Met Pro Gly Lys
 65 70 75 80
 Gly Leu Val His Lys Ser Leu Gln Thr Arg Leu Tyr Ser Asn Asp Gln
 85 90 95

WO 00/73336

PCT/SE00/01079

48

Thr Lys Ser Val Asn Val Lys Leu Leu Asn Ala Pro Gln Leu Ile Asn
 100 105 110

Val Leu Asp Pro Thr Lys Thr Ile Asp Met Glu Val Thr Leu Gly Gly
 115 120 125

Arg Ser Leu Thr Thr Thr Asn Ser Val Leu Glu Ala Lys Thr Leu Phe
 130 135 140

Pro Asp Gly Lys Thr Gly Asp Ala Ser Ala Leu Leu Asn Leu Asp Ile
 145 150 155 160

Gly Gln Lys Ala Gly Ala Ala Leu Gln Asn Leu Pro Ala Gly Glu Tyr
 165 170 175

Ser Gly Leu Val Ser Leu Val Ile Ser Gln Ala Val Thr Ala Gly
 180 185 190

<210> 4
 <211> 889
 <212> PRT
 <213> Salmonella typhi

<400> 4
 Met Tyr Tyr Leu Leu Gly Leu Cys Ser Phe Thr Ser Gln Ala Thr Leu
 1 5 10 15

Ile Pro Pro Pro Gly Phe Glu Ser Leu Leu Glu Gly Gln Thr Glu Gln
 20 25 30

Ile Glu Val Leu Leu Pro Gly His Ser Leu Gly Leu Phe Pro Val Val
 35 40 45

Val Lys Pro Asp Thr Val Gln Phe Met Ser Pro Leu Met Val Leu Glu
 50 55 60

Ser Ser Gly Leu Ala Ala Leu Pro Ala Ala Glu Arg Gln Lys Ala Leu
 65 70 75 80

Ala Ala Leu Ser Arg Pro Leu Leu Arg Asn Ser Asn Leu Val Cys Gly
 85 90 95

Val Ser Glu Ala Lys Asp Ser Ser Glu Cys Gly Tyr Val Ala Thr Asp
 100 105 110

Lys Glu Asp Val Ala Val Ile Phe Asp Glu Asn Asn Ala Gln Leu Ser
 115 120 125

Leu Phe Leu Asn Arg Asp Trp Leu Pro Asp Glu Glu Arg Arg Asp Lys
 130 135 140

Arg Trp Leu Thr Pro Thr Pro Glu Gly Val Ser Ala Phe Ile His Arg
 145 150 155 160

Gln Thr Leu Tyr Leu Ser Asp Asp Leu His Ser Arg Asn Met Thr Leu
 165 170 175

Asn Gly Ser Gly Ala Leu Gly Leu Gly Asp Gly Arg Tyr Leu Gly Gly
 180 185 190
 Asp Trp Ala Ala Ile Trp Asn Gln Ser Glu His Tyr Asn Asn Ser Gln
 195 200 205
 Ala Trp Phe Asp Asn Leu Phe Val Arg Gln Asp Leu Gly Asn Gln Tyr
 210 215 220
 Tyr Leu Gln Ala Gly Arg Met Asp Gln Arg Asn Leu Ser Ser Ala Thr
 225 230 235 240
 Gly Gly Asp Phe Gly Phe Ser Leu Leu Pro Leu Ser Arg Phe Asp Gly
 245 250 255
 Leu Arg Thr Gly Thr Thr Gln Ala Tyr Val Asn His Glu Val Asp His
 260 265 270
 Asn Ala Thr Pro Val Met Val Gln Val Thr Arg Asn Ala Arg Ile Asp
 275 280 285
 Ile Tyr Arg Gly Ser Glu Leu Leu Gly Ser Gln Phe Leu Thr Pro Gly
 290 295 300
 Met His Thr Leu Asp Thr His Ser Leu Pro Pro Gly Ser Tyr Pro Leu
 305 310 315 320
 Ala Leu Arg Val Tyr Glu Asp Gly Ile Leu Arg Arg Thr Glu Thr Gln
 325 330 335
 Pro Phe Ser Lys Gly Gly Asn Ser Phe Ser Ala Gln Thr Gln Trp Phe
 340 345 350
 Ile Gln Gly Gly Leu Glu Asp Thr Gly Asp Lys Ala Ser His Tyr Asp
 355 360 365
 Gly Glu Thr Val Met Ala Ala Gly Phe Gln Thr Gly Leu Arg Lys Asn
 370 375 380
 Ile Ser Leu Thr Glu Gly Ile Ser Leu Ala His Glu Ala Trp Tyr Ser
 385 390 395 400
 Glu Thr Arg Leu Asn Ser Gln His Ala Val Leu Asp Gly Thr Leu Asp
 405 410 415
 Leu Ser Ala Gly Ile Leu His Gly Thr Asp Ser Thr Ser Gly Asn Thr
 420 425 430
 Glu Gln Val Thr Tyr Asn Asp Gly Phe Ser Ala Ser Leu Trp Arg Asn
 435 440 445
 His Thr Glu Ser Asp Ala Cys Ser Gly Arg His Pro Gln Ser Val His
 450 455 460
 Ala Ser Met Thr Cys Gln Thr Ser Met Asn Ala Ser Leu Ser Val Ser
 465 470 475 480
 Val Gly Asn Trp Tyr Ala Leu Leu Gly Tyr Ser Thr Ser Arg Thr Glu
 485 490 495

WO 00/73336

PCT/SE00/01079

50

Gly Arg Pro Val Tyr Arg Gly Tyr Asp Asp Asn Ser Asp Lys Glu Asn
 500 505 510
 Val Phe Trp Arg Gln Ala Tyr Ile Pro Ala Ser His Arg Glu Ser Ala
 515 520 525
 Gln Ala Ser Ala Thr Tyr Ser Leu Asn Met Ala Gly Met Asn Ile Asn
 530 535 540
 Thr His Gly Gly Val Trp Arg Thr Arg Asn Asp Gly Val Asn Asp Asp
 545 550 555 560
 Gly Leu Phe Met Ser Val Ser Val Ser Tyr Ala Ser Gln Pro Pro Thr
 565 570 575
 Met Thr Gly Ser Asn Arg Tyr Thr Ser Ala Gly Thr Asp Ile His Ser
 580 585 590
 Ser Arg Asn Gln Lys Thr Gln Thr Ser Trp Asn Val Asn His Val Arg
 595 600 605
 Ser Trp Gln Gln Asp Leu Tyr Arg Glu Leu Ser Val Gly Phe Ser Gly
 610 615 620
 Tyr Asn Asp Asp Ser Trp Ser Gly Ser Leu Gly Gly Arg Met Ser Gly
 625 630 635 640
 Arg Met Gly Glu Leu Ser Ala Thr Ile Ser Asn Ser His Gln Arg Asn
 645 650 655
 Ala Gly Ser Ala Ser Ser Leu Thr Ala Gly Tyr Ser Ser Ser Leu Ala
 660 665 670
 Leu Ser Arg Asn Gly Leu Phe Trp Gly Gly Gly Gln Asp Gly Glu Pro
 675 680 685
 Ala Ser Gly Met Ala Val Asn Val Glu Ser Glu Gly Asp Glu Gly Ser
 690 695 700
 Ser Gly Lys Val Val Ser Val Arg Gly Ser Ser Gln Pro Phe Ser Leu
 705 710 715 720
 Gly Phe Gly Gln Gln Ser Leu Leu Leu Met Glu Gly Tyr Asn Ala Thr
 725 730 735
 Glu Val Thr Ile Glu Asp Ala Gly Val Ser Ser Gln Gly Met Ala Gly
 740 745 750
 Val Lys Ala Gly Gly Gly Ser Arg Cys Tyr Phe Leu Thr Pro Gly His
 755 760 765
 Leu Leu Val His Asn Ile Ser Ala Ser Met Ser Arg Leu Tyr Val Gly
 770 775 780
 Arg Val Leu Asp Lys Asp Gly Arg Pro Leu Leu Asp Ala Gln Pro Leu
 785 790 795 800
 Asn Tyr Pro Phe Leu Ser Leu Gly Pro Ser Gly Arg Phe Ser Leu Gln
 805 810 815

WO 00/73336

PCT/SE00/01079

51

Ser Glu His Lys Glu Ser Ser Leu Trp Leu Leu Ser Lys Asn Arg Ile
820 825 830

Leu Arg Cys Pro Met Ser Val His Lys Arg Arg Asp Val Met Gln Val
835 840 845

Val Gly Asp Val Arg Cys Glu Leu Ser Asp Val Asp Ala Leu Pro Gln
850 855 860

Ala Leu Gln Ile Ser Pro Arg Val Ile Arg Leu Leu Asn Val Ala Gly
865 870 875 880

Leu Leu Arg His Ser Val Gln Glu Ala
885

<210> 5

<211> 359

<212> PRT

<213> Salmonella typhi

<400> 5

Met Ser Asn Lys Met Lys Trp Thr Ser Met Thr Ala His Trp Ser Ala
1 5 10 15

Ile Ile Asn Phe Ile Arg Lys Tyr Val Tyr Pro Ala Arg Ile Ile Ala
20 25 30

Ile Leu Leu Met Ala Gly Ala Thr Leu Pro Gln Val Ala Asp Ala Ile
35 40 45

Thr Val Asp Leu Asn Tyr Asp Lys Asn Asn Val Ala Val Ile Thr Pro
50 55 60

Val Trp Ser Gln Glu Trp Ser Val Ala Asn Val Leu Gly Gly Trp Val
65 70 75 80

Cys Arg Ser Asn Arg Asn Glu Asn Glu Gly Ala Cys Glu Glu Thr His
85 90 95

Leu Val Trp Trp Tyr Ala Phe Gly Ala Tyr Ser Lys Ile Arg Leu Arg
100 105 110

Phe Arg Glu Gln Ile Ser His Ala Glu Ile Thr Leu Ile Leu Leu Gly
115 120 125

Ser Val Arg Asp Ala Cys Tyr Thr Gly Val Ile Asn Met Asn Ala Ala
130 135 140

Ala Cys Gln Trp Gly Arg Ser Leu Lys Leu Arg Ile Pro Ser Glu Glu
145 150 155 160

Leu Ala Lys Ile Pro Thr Ser Gly Thr Trp Lys Ala Thr Leu Val Leu
165 170 175

Asp Tyr Leu Gln Trp Gly Gly Asp Asp Pro Leu Gly Thr Ser Thr Thr
180 185 190

Asp Ile Thr Leu Asn Val Thr Asp His Phe Ala Glu Asn Ala Ala Ile
195 200 205

Tyr	Phe	Pro	Gln	Phe	Gly	Thr	Ala	Thr	Pro	Arg	Val	Asp	Leu	Asn	Leu
210					215					220					
His	Arg	Met	Asn	Ala	Ser	Gln	Met	Ser	Gly	Arg	Ala	Asn	Leu	Asp	Met
225					230					235					
Cys	Leu	Tyr	Asp	Gly	Gly	Val	Lys	Ala	Arg	Ser	Leu	Gln	Met	Lys	Ile
245					250					255					
Glu	Gly	Ser	Asn	Lys	Ser	Gly	Thr	Gly	Phe	Gln	Val	Ile	Lys	Ser	Asp
260					265					270					
Ser	Ala	Asp	Thr	Ile	Asp	Tyr	Ala	Val	Ser	Met	Asn	Tyr	Gly	Gly	Arg
275					280					285					
Ser	Ile	Pro	Val	Thr	Arg	Gly	Val	Glu	Phe	Ser	Leu	Asp	Asn	Val	Asp
290					295					300					
Lys	Ala	Ala	Thr	Arg	Pro	Val	Val	Leu	Pro	Gly	Gln	Arg	Gln	Ala	Val
305					310					315					
Arg	Cys	Val	Pro	Val	Pro	Leu	Thr	Leu	Thr	Thr	Gln	Pro	Phe	Asn	Ile
325					330					335					
Arg	Glu	Lys	Arg	Ser	Gly	Glu	Tyr	Gln	Gly	Thr	Leu	Thr	Val	Thr	Met
340					345					350					
Leu	Met	Gly	Thr	Gln	Thr	Pro									
355															

<210> 6

<211> 151

<212> PRT

<213> Salmonella typhi

<400> 6

Met Thr Val Tyr Leu Asp Asp Lys Asp Lys Glu Leu Leu Lys Glu Ile
1 5 10 15

Gln Lys Asp Cys Ala Gln Thr Leu Trp Gln Leu Ala Tyr Lys Val Gly
20 25 30

Leu Thr Pro Thr Pro Cys Phe Lys Arg Leu Lys Lys Leu Lys Asp Arg
35 40 45

Gly Val Ile Ile Gly Gln Phe Ala Leu Leu Asp Lys Glu Lys Leu Gly
50 55 60

Leu Ser Leu Asn Val Phe Ile Met Ile Asn Ile Ser Glu Glu Gln Tyr
65 70 75 80

Ala Ser Ile Ser Glu Lys Ile Lys Ser Met Pro Glu Val Ile Ala Phe
85 90 95

Tyr Arg Ile Ser Gly Ser Phe Asn Tyr Leu Met His Thr Val Phe Thr
100 105 110

53

Asp Met Asn Asp Tyr Tyr Ser Phe Tyr Glu Lys Ile Ile Leu Thr Asn
115 120 125

Ser Ser Ile Ser Gly Ser Ala Ser Ser Phe Val Leu Glu Gln Ile Lys
130 135 140

Glu Thr Asn Glu Leu Ser Val
145 150